

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

rotein - protein search, using sw model

n: January 6, 2004, 11:20:38 ; Search time 22 Seconds  
(without alignments)  
898.144 Million cell updates/sec

ct score: 2527

nce: 1 MGIIWVAFLEFGTSRGKEV.....STFCSDIMGNILQNLKPC 467

ng table: BLOSUM62

Gapop 10.0 , Gapext 0.5

hed: 328717 seqs, 42310858 residues

number of hits satisfying chosen parameters: 328717

um DB seq length: 0

um DB seq length: 2000000000

Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

ase : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/iaa/ECTUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/iaa/BACKFILES.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ID	Score	Query Match %	Length	DB	Description
1	2523	99.8	467	4	US-09-411-132A-3
2	1180	46.7	473	4	US-09-411-132A-7
3	1174	46.5	467	4	US-09-411-132A-9
4	1172	46.4	473	4	US-09-411-132A-8
5	1166	46.1	467	4	US-09-411-132A-10
6	1162	46.0	467	4	US-09-411-132A-4
7	1152	45.6	469	4	US-09-411-132A-5
8	1115	44.1	465	4	US-08-985-432-15
9	1115	44.1	465	4	US-09-411-132A-6
0	534	21.1	499	4	US-08-985-432-14
1	515.5	20.4	500	4	US-08-985-432-8
2	462.5	18.3	345	4	US-08-985-432-10
3	462.5	18.3	353	4	US-08-985-432-4
4	462.5	18.3	354	4	US-08-985-432-6
5	440.5	17.4	472	4	US-08-985-432-13
6	391.5	15.5	124	1	US-08-180-209B-21
7	391.5	15.5	124	1	US-08-385-745-21
8	391.5	15.5	124	3	US-08-485-388-21
9	391.5	15.5	124	3	US-08-474-853-21
0	391.5	15.5	124	4	US-09-166-205B-21
1	391.5	15.5	124	5	PCT-US94-02629-21
2	307	12.1	125	1	US-08-180-209B-20
3	307	12.1	125	1	US-08-385-745-20
4	307	12.1	125	3	US-08-485-388-20
5	307	12.1	125	3	US-08-474-853-20
6	307	12.1	125	4	US-09-166-205B-20
7	307	12.1	125	5	PCT-US94-02629-20

Sequence 64, Appl  
Sequence 27, Appl  
Sequence 27, Appl  
Sequence 27, Appl  
Sequence 27, Appl  
Sequence 27, Appl  
Sequence 27, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 19, Appl  
Sequence 16, Appl

ALIGNMENTS

RESULT 1  
US-09-411-132A-3  
Sequence 3 Application US/09411132A  
Patent No. 6558936  
GENERAL INFORMATION:  
APPLICANT: Khodadoust, Mehran  
APPLICANT: Kapeller-Libermann, Rosana  
TITLE OF INVENTION: No. 6558936 Human Lipase Proteins, Nucleic Acides Encoding  
TITLE OF INVENTION: Them, and Uses of Both of These  
FILE REFERENCE: 10147-14  
CURRENT APPLICATION NUMBER: US/09/411.132A  
CURRENT FILING DATE: 2000-09-12  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 467  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-411-132A-3

3 — see 20 No 3

Query Match	99.8%	Score	2523	DB 4	Length	467	
Best Local Similarity	99.8%	Pred. No.	1.6e-262				
Matches	466	Conservative	1	Mismatches	0	Gaps	0
QY	1	MLGIWVAFLEFGTSRGKEVYERLGC	1	MLGIWVAFLEFGTSRGKEVYERLGC	1	MLGIWVAFLEFGTSRGKEVYERLGC	1
Db	1	MLGIWVAFLEFGTSRGKEVYERLGC	1	MLGIWVAFLEFGTSRGKEVYERLGC	1	MLGIWVAFLEFGTSRGKEVYERLGC	1
QY	61	IHNPNAYOEISAVNSSSTIOASYFGT	61	IHNPNAYOEISAVNSSSTIOASYFGT	61	IHNPNAYOEISAVNSSSTIOASYFGT	61
Db	61	IHNPNAYOEISAVNSSSTIOASYFGT	61	IHNPNAYOEISAVNSSSTIOASYFGT	61	IHNPNAYOEISAVNSSSTIOASYFGT	61
QY	121	LDWNGSREYTHAVNNLRVVGAEVAF	121	LDWNGSREYTHAVNNLRVVGAEVAF	121	LDWNGSREYTHAVNNLRVVGAEVAF	121
Db	121	LDWNGSREYTHAVNNLRVVGAEVAF	121	LDWNGSREYTHAVNNLRVVGAEVAF	121	LDWNGSREYTHAVNNLRVVGAEVAF	121
QY	181	IPGLGRITGLDPAAGFFPHNTPKEV	181	IPGLGRITGLDPAAGFFPHNTPKEV	181	IPGLGRITGLDPAAGFFPHNTPKEV	181
Db	181	IPGLGRITGLDPAAGFFPHNTPKEV	181	IPGLGRITGLDPAAGFFPHNTPKEV	181	IPGLGRITGLDPAAGFFPHNTPKEV	181
QY	241	FYPNGGKMPGCCDLITPLKFN	241	FYPNGGKMPGCCDLITPLKFN	241	FYPNGGKMPGCCDLITPLKFN	241
Db	241	FYPNGGKMPGCCDLITPLKFN	241	FYPNGGKMPGCCDLITPLKFN	241	FYPNGGKMPGCCDLITPLKFN	241
QY	301	CRSYTSFKAGNCFCSKEGCGPTNGH	301	CRSYTSFKAGNCFCSKEGCGPTNGH	301	CRSYTSFKAGNCFCSKEGCGPTNGH	301
Db	301	CRSYTSFKAGNCFCSKEGCGPTNGH	301	CRSYTSFKAGNCFCSKEGCGPTNGH	301	CRSYTSFKAGNCFCSKEGCGPTNGH	301
QY	361	VKLSSSFTQGTFLRVGGAIGTGF	361	VKLSSSFTQGTFLRVGGAIGTGF	361	VKLSSSFTQGTFLRVGGAIGTGF	361
Db	361	VKLSSSFTQGTFLRVGGAIGTGF	361	VKLSSSFTQGTFLRVGGAIGTGF	361	VKLSSSFTQGTFLRVGGAIGTGF	361

361 VKLSGSEVTOGTFLRVGGAIGTGEPAIVSGKLEPQMTYTKLIDAEVAVNGNITSVQFTW 420  
421 KKHLPEDSQKLGAEWINTSGKYGKSTFCSDIMGNILQNLKPC 467  
421 KKHLPEDSQKLGAEWINTSGKYGKSTFCSDIMGNILQNLKPC 467

AT 2  
9-411-132A-7  
Sequence 7, Application US/09411132A  
Patent No. 6558936  
GENERAL INFORMATION:  
APPLICANT: Khodadoust, Mehran  
PLICANT: Kapeller-Libermann, Rosana  
TITLE OF INVENTION: No. 6558936el Human Lipase Proteins, Nucleic Acides Encoding  
TITLE OF INVENTION: Them, and Uses of Both of These  
CURRENT APPLICATION NUMBER: US/09/411,132A  
CURRENT FILING DATE: 2000-09-12  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.1  
LENGTH: 467  
TYPE: PRT  
ORGANISM: Mus musculus  
9-411-132A-7

Query Match 46.7%; Score 1180; DB 4; Length 473;  
Best Local Similarity 49.0%; Pred. No. 5.4e-118;  
Matches 231; Conservative 73; Mismatches 159; Indels 8; Gaps 6;

QY 1 MLCIWIIVAFPFSTGRGKVCYERLGCFCFQDGLPWTFTSTELVGLPWSPEKINTRELLYT 60  
DB 1 MLCIWIIVAFPFSTGRGKVCYERLGCFCFQDGLPWTFTSTELVGLPWSPEKINTRELLYT 60  
QY 61 IHNPNAYQEISAVNSSTIQASYFGTDKITRINIAGW--KTDGKWQDMCNVLLQLEEDINC 118  
DB 61 IHNPNAYQEISAVNSSTIQASYFGTDKITRINIAGW--KTDGKWQDMCNVLLQLEEDINC 120  
QY 119 INLDWINGSR-EYIHAVNNLRVVGAEVAFIDVLMKKFESPSKVHLIGHSLGAHLAGEA 177  
DB 121 ICVDWKKGSQTSYTOAANNRVVGAQVQMLSMLSANYSYSPSQVQLIGHSLGAHVAGEA 180  
QY 178 GSRIPLGLGRITGLDPAGPFFHNTPKVEVRLDPSDANFVDVHTNAARILFELGVGTIDACG 237  
DB 181 GSRTPLGLGRITGLDPVEANFEGTPEEVLDPDADFVDVHTDAAPLIPFLFGFTSQMG 240  
QY 238 HLDYPNGGKHPGCC-EDLITPLKFNFNAYKKEWASFPDCNHARSYQFYAESILNPDFA 296  
DB 241 HFDFFPNGGQMPGCKKNALSQIV--DIDGWSGTRDFVACNHLRSYKYLESILNPDGF 298  
QY 297 IAYPCRSYTSFKAGNCFECSKEGCTMGHFAHDFHFKMKMGSHYPLNTGSLSPFARWR 356  
DB 299 AAYPCASYRPFESNKCFCPCDQCPQMGHYADKFAVKT-SDETQYFELNTGDSNPFARWR 357  
QY 357 HKLSVKLSGSEVTOGTFLRVGGAIGTGEPAIVSGKLEPQMTYTKLIDAEVAVNGNITSV 416  
DB 358 YRVSLTFSGTFTV-GQVKVSLFGSNGTRQCDIFRGIIKPGATHSNEFDKLDVGTIEKV 416  
QY 417 QFTWKHLPEDSQKLGAEWINTSGKYGKSTFCSDIMGNILQNLKPC 467  
DB 417 KFLNNHVNVPFPKVGAAKITVQKGEERTENHFCSEETVREDILLTLTPC 467

AT 3  
9-411-132A-9  
Sequence 9, Application US/09411132A  
Patent No. 6558936  
GENERAL INFORMATION:  
APPLICANT: Khodadoust, Mehran  
PLICANT: Kapeller-Libermann, Rosana  
TITLE OF INVENTION: No. 6558936el Human Lipase Proteins, Nucleic Acides Encoding  
TITLE OF INVENTION: Them, and Uses of Both of These

FILE REFERENCE: 10147-14  
CURRENT APPLICATION NUMBER: US/09/411,132A  
CURRENT FILING DATE: 2000-09-12  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 467  
TYPE: PRT  
ORGANISM: Canis familiaris  
US-09-411-132A-9

Query Match 46.5%; Score 1174; DB 4; Length 467;  
Best Local Similarity 48.2%; Pred. No. 2.3e-117;  
Matches 227; Conservative 78; Mismatches 158; Indels 8; Gaps 6;

QY 1 MLCIWIIVAFPFSTGRGKVCYERLGCFCFQDGLPWTFTSTELVGLPWSPEKINTRELLYT 60  
DB 1 MLCIWIIVAFPFSTGRGKVCYERLGCFCFQDGLPWTFTSTELVGLPWSPEKINTRELLYT 60  
QY 61 IHNPNAYQEISAVNSSTIQASYFGTDKITRINIAGW--KTDGKWQDMCNVLLQLEEDINC 118  
DB 61 IHNPNAYQEISAVNSSTIQASYFGTDKITRINIAGW--KTDGKWQDMCNVLLQLEEDINC 120  
QY 119 INLDWINGSR-EYIHAVNNLRVVGAEVAFIDVLMKKFESPSKVHLIGHSLGAHLAGEA 177  
DB 121 ICVDWKKGSQTSYTOAANNRVVGAQVQMLSMLSANYSYSPSQVQLIGHSLGAHVAGEA 180  
QY 178 GSRIPLGLGRITGLDPAGPFFHNTPKVEVRLDPSDANFVDVHTNAARILFELGVGTIDACG 237  
DB 181 GSRTPLGLGRITGLDPVEANFEGTPEEVLDPDADFVDVHTDAAPLIPFLFGFTSQMG 240  
QY 238 HLDYPNGGKHPGCC-EDLITPLKFNFNAYKKEWASFPDCNHARSYQFYAESILNPDFA 296  
DB 241 HLDFFPNGGQMPGCKKNALSQIV--NLDGIWEGTRDFVACNHLRSYKYSESILNPDGF 298  
QY 297 IAYPCRSYTSFKAGNCFECSKEGCTMGHFAHDFHFKMKMGSHYPLNTGSLSPFARWR 356  
DB 299 AAYPCASYRPFESNKCFCPCDQCPQMGHYADKFAVKT-SDETQYFELNTGDSNPFARWR 357  
QY 357 HKLSVKLSGSEVTOGTFLRVGGAIGTGEPAIVSGKLEPQMTYTKLIDAEVAVNGNITSV 416  
DB 358 YGVSLTFSGRAT-QQAKVALFGSKGNTHQFNFKGILKPGSTHSNEFDKLDVGTIEKV 416  
QY 417 QFTWKHLPEDSQKLGAEWINTSGKYGKSTFCSDIMGNILQNLKPC 467  
DB 417 KFLNNHVNVPFPKVGAAKITVQKGEERTVHVSFSESTVREDVLLTLTPC 467

RESULT 4  
US-09-411-132A-8  
Sequence 8, Application US/09411132A  
Patent No. 6558936  
GENERAL INFORMATION:  
APPLICANT: Khodadoust, Mehran  
PLICANT: Kapeller-Libermann, Rosana  
TITLE OF INVENTION: No. 6558936el Human Lipase Proteins, Nucleic Acides Encoding  
TITLE OF INVENTION: Them, and Uses of Both of These  
CURRENT APPLICATION NUMBER: US/09/411,132A  
CURRENT FILING DATE: 2000-09-12  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 8  
LENGTH: 473  
TYPE: PRT  
ORGANISM: Rattus norvegicus  
US-09-411-132A-8

Query Match 46.4%; Score 1172; DB 4; Length 473;  
Best Local Similarity 49.3%; Pred. No. 3.9e-117;  
Matches 232; Conservative 73; Mismatches 158; Indels 8; Gaps 6;

QY 1 MLCIWIIVAFPFSTGRGKVCYERLGCFCFQDGLPWTFTSTELVGLPWSPEKINTRELLYT 60

```
1 MLTLVSLFLGAAQGEVYDNLGCFSDAEPWAGTAIRPLKLLPWSPEKINTRELLYT 60
61 IHNPNAYQEIISAVNSSTIQASYFGTDKTRINIAW--KTDGKWORDMKNVLLQLEDINC 118
61 NENPTAFQTLQSLDPLTIGASNFQVARKTRFIHGFIDKGEENVVDMCKMFMQVEVNC 120
119 INLDWINGSR-EYIHAVNNLRVVGAEVAFIDVLMKKFEYSPSKVHLIGHSLGAHLAGEA 177
121 ICVDWKKGSQITTYQOANNVRVVGAAQVAMQIDILVXNYSYSPSKVHLIGHSLGAHVAGEA 180
178 GSRIPGLGRITGLDPAGFPFHTPKVELRDPDSANFVDVHTNAARILFELGVGTIDACG 237
181 GSRTPGLGRITGLDPVEANFETGTEEVNRLDSDADFVDVHTDAAPLIPFLGFGTQMQSG 240
238 HLDFFPNGGKHPGC-EDLITPLKKNFNAYKEMASFFDCNHNARSYQFYAESILNPDAF 296
241 HLDFFPNGGQMPGCKKNALSQIV--DIDGWSGTRDFVACNHLRSYKYLESILNPDPGF 298
297 IAYPCRSYTSFKAGNCFECSKEGCPMGHFAHDFHFKNMKTNGSHYFLNTGSLSPFARWR 356
299 AAYPCASYKDFESNCKFPDQCPQMGHYADKFAKGS-GDEPKQFELNTGEAKNFARWR 357
357 HKLSVKLSGSEVTOGTVELRVGGAIGKTGEFAIVSGKLEPMTYTKLIDADVNGNITSV 416
358 YRVSILLSGRWVT-GQKVALFGSKGNTRQVDIFRGLIKPGATHSSEFADKLDVGTIEKV 416
417 QFIWKHKLFPESQNKLGAEWINTSGKYGYKSTFCSDQDIMGPNILQNLKPC 467
417 KFLMNNQVINFSPFRVGAAKITVQKGEERTYFNFCSEETVREDTLLTLTPC 467
5
-411-132A-10
ence 10, Application US/09411132A
nt No. 6558936
RAL INFORMATION:
LICANT: Khodadoust, Mehran
LICANT: Kapeller-Libermann, Rosana
LE OF INVENTION: No. 6558936el Human Lipase Proteins, Nucleic Acides Encoding
LE OF INVENTION: Them, and Uses of Both of These
E REFERENCE: 10147-14
RENT APPLICATION NUMBER: US/09/411,132A
ER OF FILING DATE: 2000-09-12
BER OF SEQ ID NOS: 10
TWARE: PatentIn Ver. 2.1
ID NO 10
NGTH: 467
PE: PRT
GANISM: Canis familiaris
-411-132A-10
Y Match 46.1%; Score 1166; DB 4; Length 467;
Local Similarity 48.0%; Pred. No. 1.7e-116;
hes 226; Conservative 79; Mismatches 158; Indels 8; Gaps 6;
1 MLGIWIVAFLEFGTSRGEVCYERLGCGLPWTRFTSTELVGLPWSPEKINTRELLYT 60
1 MVSINTIALFLGAAKAKEVCYEOIGCFSDAEPWAGTAIRPLKVLWSPERIGTRFLYT 60
61 IHNPNAYQEIISAVNSSTIQASYFGTDKTRINIAW--KTDGKWORDMKNVLLQLEDINC 118
61 NKNPNFOTLLPSPDSTIEASNFQTKKTRFIHGFIDKGEENVLLDMCKMFKVEVNC 120
119 INLDWINGSR-EYIHAVNNLRVVGAEVAFIDVLMKKFEYSPSKVHLIGHSLGAHLAGEA 177
121 ICVDWKKGSQISYQOANNVRVVGAAQVAMLSMANSYSYSPQVQLIGHSLGAHVAGEA 180
178 GSRIPGLGRITGLDPAGFPFHTPKVELRDPDSANFVDVHTNAARILFELGVGTIDACG 237
181 GSRTPGLGRITGLDPVEASFQGTPEEVRDLPTDADFVDVHTDAAPLIPFLGFGTQMQSG 240
238 HLDFFPNGGKHPGC-EDLITPLKKNFNAYKEMASFFDCNHNARSYQFYAESILNPDAF 296
```

```
Db 241 HLDFFPNGGEMPGCKKNALSQIV--DLDGIEWGTRDFVACNHLRSYKYSESILNPDPGF 298
QY 297 IAYPCRSYTSFKAGNCFECSKEGCPMGHFAHDFHFKNMKTNGSHYFLNTGSLSPFARWR 356
Db 299 ASYPCASYRAFESNCKFPDQCPQMGHYADKFAVKT-SDETQKIFLNTGSSNFARWR 357
QY 357 HKLSVKLSGSEVTOGTVELRVGGAIGKTGEFAIVSGKLEPMTYTKLIDADVNGNITSV 416
Db 358 YGVSIITLSGKRAT-GQAKVALFGSKGNTHQFNIFKGLKPGSTHSEFADKLDVGTIEKV 416
QY 417 QFIWKHKLFPESQNKLGAEWINTSGKYGYKSTFCSDQDIMGPNILQNLKPC 467
Db 417 KFLMNNVNVNPTFPKVGAAKITVQKGEKTVHVSFCSESTVREDVLLTLTPC 467
RESULT 6
US-09-411-132A-4
; Sequence 4, Application US/09411132A
; Patent No. 6558936
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: No. 6558936el Human Lipase Proteins, Nucleic Acides Encoding
; TITLE OF INVENTION: Them, and Uses of Both of These
; FILE REFERENCE: 10147-14
; CURRENT APPLICATION NUMBER: US/09/411,132A
; CURRENT FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-411-132A-4
Query Match 46.0%; Score 1162; DB 4; Length 467;
Best Local Similarity 48.6%; Pred. No. 4.6e-116;
Matches 229; Conservative 75; Mismatches 159; Indels 8; Gaps 6;
QY 1 MLGIWIVAFLEFGTSRGEVCYERLGCGLPWTRFTSTELVGLPWSPEKINTRELLYT 60
Db 1 MLIFWTITLFLGAAKAKEVCYEDLGCFSDFPQMGTAIRPLKILPWSPEKIGTRFLYT 60
QY 61 IHNPNAYQEIISAVNSSTIQASYFGTDKTRINIAW--KTDGKWORDMKNVLLQLEDINC 118
Db 61 NENPNFQIILLSDPSTIEASNFQMDKTRFIHGFIDKGSWVTDCKLFEVEVNC 120
QY 119 INLDWINGSR-EYIHAVNNLRVVGAEVAFIDVLMKKFEYSPSKVHLIGHSLGAHLAGEA 177
Db 121 ICVDWKKGSQATYQOANNVRVVGAAQVAMQIDILTEYSYPPSKVHLIGHSLGAHVAGEA 180
QY 178 GSRTPGLGRITGLDPAGFPFHTPKVELRDPDSANFVDVHTNAARILFELGVGTIDACG 237
Db 181 GSKTPEGLSRITGLDPVEASFESTPEEVRDLDPDSADFDVHTDAAPLIPFLGFGTQMQMG 240
QY 238 HLDFFPNGGKHPGC-EDLITPLKKNFNAYKEMASFFDCNHNARSYQFYAESILNPDAF 296
Db 241 HLDFFPNGGSEMPGCKKNALSQIV--DLDGIEWGTRDFVACNHLRSYKYLESILNPDPGF 298
QY 297 IAYPCRSYTSFKAGNCFECSKEGCPMGHFAHDFHFKNMKTNGSHYFLNTGSLSPFARWR 356
Db 299 AAYPCTSYKSFESDKCFPCPDQCPQMGHYADKFAKRTSEEQ-QKFPFLNTGEASNARWR 357
QY 357 HKLSVKLSGSEVTOGTVELRVGGAIGKTGEFAIVSGKLEPMTYTKLIDADVNGNITSV 416
Db 358 YGVSIITLSGTRAT-GQKVALFGKNKNTHQISLFRGILRPGSTHSEFADKLDVGTIEKV 416
QY 417 QFIWKHKLFPESQNKLGAEWINTSGKYGYKSTFCSDQDIMGPNILQNLKPC 467
Db 417 KFLMNNVNVNPTLPKVGATKITVQKGEKTVNFCSESTVREDTLLTLTPC 467
```

102 7  
 103 -411-132A-5  
 104 nence 5, Application US/09411132A  
 105 ent No. 6558936  
 106 ERAL INFORMATION:  
 107 APPLICANT: Khodadoust, Mehran  
 108 APPLICANT: Kapeller-Libermann, Rosana  
 109 TITLE OF INVENTION: NO. 6558936el Human Lipase Proteins, Nucleic Acides Encoding  
 110 TLE OF INVENTION: Them, and Uses of Both of These  
 111 LE REFERENCE: 10147-14  
 112 RRENT APPLICATION NUMBER: US/09/411,132A  
 113 RRENT FILING DATE: 2000-09-12  
 114 MBER OF SEQ ID NOS: 10  
 115 FTWARE: Patentin Ver. 2.1  
 116 ID NO 5  
 117 LENGTH: 469  
 118 TYPE: PRT  
 119 ORGANISM: Homo sapiens  
 120 -411-132A-5  
 121 ry Match 45.6%; Score 1152; DB 4; Length 469;  
 122 t Local Similarity 47.8%; Pred. No. 5.5e-115;  
 123 ches 225; Conservative 68; Mismatches 172; Indels 6; Gaps 5;  
 124  
 125 1 MLCGIWIVAFLEFGTSRGKVCYERLCGKDPGLPWTSTELVGLPWSPEKINTREFLLYT 60  
 126 1 MLPPWILGLLLATVKGKVCYGLGCFSDERPWAGTLQRPVKLLPWSPEIDIDTRFLLYT 60  
 127  
 128 61 IHNPNAYQEISAVNSSTIOASVFGTDKITRINIAGW--KTDGKWORDMNCVLLQLEDINC 118  
 129 61 NENPNPQILITGTEPDTIEASNPQLDRKTRFIHGFIDKAEDESPDMCKMFEKYNVC 120  
 130 119 INLDWINGSRE-YIHAVNLRVVGAEVAFIDVLMKKFEYSKVLHIGSLGAHLAGEA 177  
 131 121 ICVDWHRGSRAMYTQAVQIRVVGAETFLIOALSTQLGYSLEDVHVHIGSLGARTAAEA 180  
 132 178 GSRIPG-LGRITGLDPAGFPFHNTPREVELDPSDFANFVDVHTNAARILFELGVGTIDAC 236  
 133 181 GRLGRVGRITGLDPAGFCQDEPEVELDPSDAFVDVHTDSSPIVPSLFGMSQKV 240  
 134 237 GHLDFFPNGKHEMPCGEDLITPLKKNFNAYKEMASFFDCNHARSYQFYAESILNPDFA 296  
 135 241 GHLDFFPNGKHEMPCGKKNVLTIT--DIDG1WEGIGGFVSCNHLSEFYEYSVSLNPDGF 299  
 136 297 IAYPCASYTSFKAGNCFKCEGCPMGMHAFDRPHFNKMTGSHYFLNTGSLSPFAWR 356  
 137 300 LGYPCASYDEFQESKCFPCPAEGCPKMGHYAQDFKGTSAVEQT--FFLNTGESGNFTSWR 358  
 138 357 HKLSVKLSGSEVTOGTVFLRVGGAIGKTGEPAIVSGKLEPGMTYTKLIDADVNGNITSV 416  
 139 359 YKVSVTLGSKGKNGYIRIALYGSNENSKQYBIFKGLKPDASH\*CAIDVDVFNVGKIQV 418  
 140 417 QFTWKXHLFPEDSONKLGAEWVINTSGKYKSTFCSDQIMGNILQNLKPC 467  
 141 419 KFLWNRGINLSPKLGASQITVQSGEDGTEYFNFCSSDTVEENVLQSLYPC 469  
 142  
 143 -985-492-15  
 144 nence 15, Application US/08985492  
 145 ent No. 639530  
 146 ERAL INFORMATION:  
 147 APPLICANT: Jaye, Michael C.  
 148 APPLICANT: Doan, Kim-Anh T.  
 149 APPLICANT: Krawiec, John A.  
 150 APPLICANT: Lutch, Kevin J.  
 151 APPLICANT: Amin, Dilip V.  
 152 APPLICANT: South, Victoria J.  
 153 TITLE OF INVENTION: LLG POLYPEPTIDES OF THE TRIACYLGLYCEROL  
 154 TITLE OF INVENTION: LIPASE FAMILY, AND COMPOSITIONS AND METHODS FOR THEIR USE  
 155 TITLE OF INVENTION: IN ENZYMIC HYDROLYSIS, AND PROTEIN AND GENE THERAPIES  
 156 NUMBER OF SEQUENCES: 31  
 157 CORRESPONDENCE ADDRESS:

102 103  
 104 ADDRESS: Rhone-Poulenc Rorer Inc.  
 105 STREET: 500 Arcola Rd. 3C43  
 106 CITY: Collegeville  
 107 STATE: PA  
 108 COUNTRY: USA  
 109 ZIP: 19426  
 110 COMPUTER READABLE FORM:  
 111 MEDIUM TYPE: Floppy disk  
 112 COMPUTER: IBM PC compatible  
 113 OPERATING SYSTEM: PC-DOS/MS-DOS  
 114 SOFTWARE: Patentin Release #1.0, Version #1.30  
 115 CURRENT APPLICATION DATA:  
 116 APPLICATION NUMBER: US/08/985,492  
 117 FILING DATE:  
 118 CLASSIFICATION: 800  
 119 ATTORNEY/AGENT INFORMATION:  
 120 NAME: Fehlner Ph.D., Paul F.  
 121 REGISTRATION NUMBER: 35,135  
 122 REFERENCE/DOCKET NUMBER: A2582-US  
 123 TELEPHONE: (610)454-3839  
 124 TELEFAX: (610)454-3808  
 125 INFORMATION FOR SEQ ID NO: 15:  
 126 SEQUENCE CHARACTERISTICS:  
 127 LENGTH: 465 amino acids  
 128 TYPE: amino acid  
 129 STRANDEDNESS:  
 130 TOPOLOGY: linear  
 131 MOLECULE TYPE: protein  
 132 US-08-985-492-15  
 133  
 134 Query Match 44.1%; Score 1115; DB 4; Length 465;  
 135 Best Local Similarity 47.8%; Pred. No. 5.2e-111;  
 136 Matches 226; Conservative 73; Mismatches 160; Indels 14; Gaps 10;  
 137  
 138 QY 1 MLCGIWIVAFLEFGTSRGKVCYERLCGKDPGLPWTSTELVGLPWSPEKINTREFLLYT 60  
 139 Db 1 MLPPWILGLLLATVKGKVCYGLGCFSDERPWAGTLQRPVKLLPWSPEIDIDTRFLLYT 59  
 140 QY 61 IHNPNAYQEISAVNSSTIOASVFGTDKITRINIAGW--KTDGKWORDMNCVLLQLEDINC 118  
 141 Db 60 NENPNPQILITGTEPDTIEASNPQLDRKTRFIHGFIDKAEDESPDMCKMFEKYNVC 118  
 142 QY 119 INLDWINGSRE-YIHAVNLRVVGAEVAFIDVLMKKFEYSKVLHIGSLGAHLAGEA 177  
 143 Db 119 ICVDWHRGSRITGTCASQIRIIVGAEVAFVEFLQSAFGYSPSNVHVHIGSLGAHLAGEA 178  
 144 QY 178 GSRIPG-LGRITGLDPAGFPFHNTPREVELDPSDFANFVDVHTNAARILFELGVGTIDAC 236  
 145 Db 179 GRLGRVGRITGLDPAGFCQDEPEVELDPSDAFVDVHTDSSPIVPSLFGMSQKV 238  
 146 QY 237 GHLDFFPNGKHEMPCGEDLITPLKKNFNAYKEMASFFDCNHARSYQFYAESILNPDFA 296  
 147 Db 239 GHLDFFPNGKHEMPCGKKNVLTIT--DIDG1WEGITRDPAAACNHLRSKYKTYTDSIVNPDGF 297  
 148 QY 297 IAYPCASYTSFKAGNCFKCEGCPMGMHAFDRPHFNKMTN--GSHYFLNTGSLSPFA 354  
 149 Db 298 AGFPCASYNVFTANKPCPCGCGCPMGHYADRY---PGKTNDVGQKCYLCTGDASNFA 354  
 150 QY 355 WHKLSVKLSGSEVTOGTVFLRVGGAIGKTGEPAIVSGKLEPGMTYTKLIDADVNGNIT 414  
 151 Db 355 WRYKVSVTLGSKKVT-GHILVSLFGNKNKQYBIFKGLKPDASH\*CAIDVDVFNVGDLQ 413  
 152 QY 415 SYQFIWKXHLFPEDSONKLGAEWVINTSGKYKSTFCSDQIMGNILQNLKPC 467  
 153 Db 414 MYKFIWNNVINPTLPRVGASKII-VETNVGKQFNFCSPETVREEVLLTLTPC 465  
 154  
 155 RESULT 9  
 156 US-09-411-132A-6  
 157 ; Sequence 6, Application US/09411132A  
 158 ; Patent No. 6558936  
 159 ; GENERAL INFORMATION:



PPLICANT: Khodadoust, Mehran  
PPLICATION: Kapeller-Liebermann, Rosana  
TILE OF INVENTION: NO. 655936el Human Lipase Proteins, Nucleic Acids Encoding  
TITLE OF INVENTION: Them, and Uses of Both of These  
LE REFERENCE: 10147-14  
RRENT APPLICATION NUMBER: US/09/411,132A  
RRENT FILING DATE: 2000-09-12  
MBER OF SEQ ID NOS: 10  
FWARE: Patentin ver. 2.1

ID NO 6  
ENGTH: 465  
YPE: PRT  
RGANISM: Homo sapiens  
-411-132A-6

ry Match      44.1%; Score 1115; DB 4; Length 465;  
Local Similarity 47.8%; Pred. No. 5.2e-111;  
ches 226; Conservative 73; Mismatches 160; Indels 14; Gaps 10;

1 MLGIWIVAFLLFGTSRGKEVCYERLGCPCDGLPWTRTFSTELVCLPWSPEKTNPLFLYT 60  
||| ||| : : | : | : | : | : | : | : | : | : | : | : | : |  
1 MPLWTLS-LILGAVAGKEVCYERLGCSDDSPSGITERPLPHLPWSPDVNPLFLT 59  
||| ||| : : | : | : | : | : | : | : | : | : | : | : | : |  
  
61 IHNPNAYQBIISAVNSSTIQASVFGTDKITRIINIAGW--KTGKWQRDMCNVLQLLEDINC 118  
||| ||| : : | : | : | : | : | : | : | : | : | : | : | : |  
60 NEHPNNFQEVAA-DSSSISGSNFKNRKTRFIHGFIDKGEEENLANVCNLKFVESVNC 118  
||| ||| : : | : | : | : | : | : | : | : | : | : | : | : |  
  
119 INLDWMINGSE-YIHAVNNLRVGAEYAFIDVLMKFYSPSKVHLIGHSLGAHLAGEA 177  
||| ||| : : | : | : | : | : | : | : | : | : | : | : | : |  
119 ICVDMWGKSRTGYTOASNIRIVGAEYAFVEFFQSAGFYSPSNHVHIGHSIGAHAAGEA 178  
||| ||| : : | : | : | : | : | : | : | : | : | : | : | : |  
  
178 GSRIPO-LGRITGLDPAGFPFHNPKEVRLDPPSDANFDVDVIHTNAARILFELGVGTIDAC 236  
||| ||| : : | : | : | : | : | : | : | : | : | : | : | : |  
179 GRRTNGTCRITGLDPAEPCCQGIPELVRLDPSDAKFVDVIHTDGAPIVPNFLGFGSQVV 238  
||| ||| : : | : | : | : | : | : | : | : | : | : | : | : |  
  
237 GHLDFFPNGCKMPGCCDLIPLIKFNPNAYKKEMASFFDCNHARSYQYASSLINPADP 296  
||| ||| : : | : | : | : | : | : | : | : | : | : | : | : |  
239 GHLDFFPNGVGVMPCGCKNILSQI-VDIGTWEGTRDFAACNHLRSKYTYTDSIVNPDGF 297  
||| ||| : : | : | : | : | : | : | : | : | : | : | : | : |  
  
297 IAPPCRSYTSFKAGMCFCEKGECPMGCHFPADRFHEFXNMKTN--GSHYFLMTGSLSPAR 354  
||| ||| : : | : | : | : | : | : | : | : | : | : | : | : |  
298 AGFFCASYNVTANKFCFFCPSGGCPQMGHYADRY---PGKTVGQKPFYDTODDASNFAR 354  
||| ||| : : | : | : | : | : | : | : | : | : | : | : | : |  
  
355 WRHKJSVKLSGSSEVTQGVTVRVUGAIGKTGEFAIVSGBKLEPGMYTKLIDADVNVGNIT 414  
||| ||| : : | : | : | : | : | : | : | : | : | : | : | : |  
355 WRYKSVTLTGKKVT-GHILVSIFGNKENSQYEIPKGTCLKPDSTHSNEFSDVDVDGDQ 413  
||| ||| : : | : | : | : | : | : | : | : | : | : | : | : |  
  
415 SVQFTWKXHLFEDSKNLGAENVINTSGKCYGSTPCSQDIMGFNILONKBC 467  
||| ||| : : | : | : | : | : | : | : | : | : | : | : | : |  
414 MKVFYNNVINPTLRVYGASKII-VETNVGQPFCSPETVREEVLLTLTPC 465  
||| ||| : : | : | : | : | : | : | : | : | : | : | : | : |

r 10  
-985-492-14  
ence 14, Application US/08985492  
nt No. 639530

NAL INFORMATION:  
PPLICANT: Jaye, Michael C.  
PPLICANT: Doan, Kim-Anh T.  
PPLICANT: Kwawiec, John A.  
PPLICANT: Lynch, Kevin J.  
PPLICANT: Amin, Dilip V.  
PPLICANT: South, Victoria J.

TITLE OF INVENTION: LLG POLYPETIDES OF THE TRIACYLYLCEROL  
TITLE OF INVENTION: LIPASE FAMILY, AND COMPOSITIONS AND METHODS FOR THEIR USE  
TITLE OF INVENTION: IN ENZYMATIC HYDROLYSIS, AND PROTEIN AND GENE THERAPIES  
UMBER OF SEQUENCES: 31

ORRESPONDENCE ADDRESS:  
ADDRESS: Rhone-Poulenc Rorer Inc.  
REET: 500 Arcola Rd. 3c43  
ITY: Collegeville  
ATE: PA  
OUNTRY: USA

```

; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent'n Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,492
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Fehlner Ph.D., Paul F.
; REGISTRATION NUMBER: 35,135
; REFERENCE/DOCKET NUMBER: A2582-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 499 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-985-492-14

Query Match 21.1%; Score 534; DB 4; Length 499;
Best Local Similarity 33.3%; Pred. No. 28-48;
Matches 145; Conservative 64; Mismatches 183; Indels 44; Gaps 14;

QY 51 KINTRFLYLTHNPAYQEISAVNSSTIOASVFGTDKITRINAGWKTDGK-----WORD 105
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 47 EMKTRFLFGETNGCC--QIRINHPDTLOECGFNSSLPLVMIHGSVDGVLENMIQMW 104
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 106 MCNVLLQLEDINCINLDWINGSRE-YIHAVNLRVVGAEVAFIDVLMKKEYSPSKVHL 164
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 105 AALKSQPAQPVNVLVDWITLADHDHTIIVRNLVGVKEVAALURWLEESVQLSRSHVHL 164
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 165 IGHSIGLAHLAGEAGSRIPG---LGRITGLDPAGPFFPHNTPKREVRLDPSDANFVDVIHTNA 221
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 165 IGVSLGARHVSFGAGSSIGGTHKIGRITGLDAAGPLFEGSAPSNRLSPDANFVDAIHT-F 223
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 222 ARILFEIGVGTIDACGLDYPYNGGXHMGCE--DLITPLLLKFNFNAYKKEMASFFDCNH 279
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 224 TREHMGLSVGIKQPIGHYDFYPNGGSPFGCFHFLYRHHIAQHGFNA----ITQTIKCSH 279
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 280 ARSYQFAESTLNPDA-FLIAPCRSVTSFKAGNCFECSKGCPTMGHFAFRHFKNKT 337
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 280 ERSVHLFIDSLIHAGTQSMAYPCDWNVSQGLCLSCKKGRCNTLG-----YHVRQSPRS 334
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 338 NGSHYFLNTGSLGPFARWRHKLVSKLGSSEVT--QGTVFLRVGGAGIKGTGEFAIVSGK-L 394
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 335 KSKRLFLVTRAQSPFKVYHQLKIQINOTETPIQTTFMTSLGTTKKNQKIPITLCKGI 394
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 395 EPGWTYTKLIDADVNGNITSVFQIWK-----KHLFEDSQNKLGAEMVINT-- 440
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 395 ASNTYGFLLTLDVIGELIMIKFWNSAVWNVDTVQTIIPWSTGPRHSGLVLTIR 454
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 441 --SGKYGKSTFCSD 454
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 455 VKAGETQQRMTFCSEN 470
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 11
US-08-985-492-8
; Sequence 8, Application US/08985492
; Patent No. 6395530
; GENERAL INFORMATION:
; APPLICANT: Jaye, Michael C.
; APPLICANT: Doan, Kim-Anh T.
; APPLICANT: Krawiec, John A.
; APPLICANT: Lynch, Kevin J.
; APPLICANT: Amin, Dilip V.

```

RESULT 11

RESOLUTION  
US-08-985-492-8  
; Sequence 8, Application US/08985492

FACE NO. 0393330  
: GENERAL INFORMATION:

;  
; GENERAL INFORMATION:  
; APPLICANT: Jaye, Michael C.  
; APPLICANT: Doan, Kim-Anh T.  
; APPLICANT: Krawiec, John A.  
; APPLICANT: Lynch, Kevin J.  
; APPLICANT: Amin, Dilip V.

APPLICANT: South, Victoria J.  
 TITLE OF INVENTION: LIG POLYPEPTIDES OF THE TRIACYLGLYCEROL  
 TITLE OF INVENTION: LIPASE FAMILY, AND COMPOSITIONS AND METHODS FOR THEIR USE  
 TITLE OF INVENTION: IN ENZYMATIC HYDROLYSIS, AND PROTEIN AND GENE THERAPIES  
 NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Rhone-Poulenc Rorer Inc.  
 STREET: 500 Arcola Rd. 3C43

CITY: Collegeville

STATE: PA

COUNTRY: USA

ZIP: 19426

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/985,492

FILING DATE:

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Fehner Ph.D., Paul F.

REGISTRATION NUMBER: 35,135

REFERENCE/DOCKET NUMBER: A2582-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (610)454-3839

TELEFAX: (610)454-3808

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 500 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

985-492-8

ary Match 20.4%; Score 515.5; DB 4; Length 500;  
 at Local Similarity 31.2%; Pred. No. 1.9e-46;  
 :ches 152; Conservative 72; Mismatches 178; Indels 85; Gaps 22;

20 VCYERLCGKDFGKPLWTRFTSTELVGLPWSPE-----KINTRELLYTIH 62

13 LCY-----CFAAGSP-----VFPGEGRLEDKHKPKATQTEVKPSVRNLRATSK 57

63 NP---NAYQEI SAVNSTIQASYFGTDKITRINAGWKTDG---KWQDMCNVLLQLE-D 115

58 DPEHEGCV--LSVGHSGPLEDCSFNWTAKTFIIGHWTMSGIFENWJHKLVSALHTREKD 115

116 INCINLDWINGSRE-YIHAVNNLRVVGAEVAYFIDVLMKKFEYSPSKVHLIGHSLGAHLA 174

116 ANVVVDWDLPLAHLQLYTDAVNNTRVVGHSIARMLDWLQEKDQDFSLGNVHLIGYSLGAHVA 175

175 GEAGSRIPG-LGRITGLDPAGPFPHNTPEKVRLDPSDANFVDVIHTNAARILFELGVGTI 233

176 GYAGNFVKGTGVRITGLDPAGPFEGADIKRLSPDDADFDVLHTYTRS--FGLSIGIQ 233

234 DACGHLDFYPNGKHPGCG--EDLITPLKFNFNAYKEMASFFDCNHARSYQFYAESIL 291

234 MPVGHIDIYPNGDFQPGGLNDVLGSI-----AY-GTITEVVKCEHERAVHLFVDSL 286

292 NPD-AFTAYPCRSYTSKAGNCFCSKEGCPVGHFADPHFKNMKTNGSHVFLNTGSL 350

287 NQDKPSFAQCTDSNRFKGCICLSCKRNCNSIGYNAKQVRNK---RNSKMYLXTRAGM 342

351 PFAAWR-----HKL SVKLSGSEVTQGVFLRVGG--AIGKTGEFAVSGKLPFGMTYTKL 403

343 PFRVYHYQMKIHVFSYKNWGE--IETFFVYLTGTNADSCQTLPLEIVE-RISQATNFTL 399

404 IDADVNTGNTITSVQFIWK-----KHLFEDSQNKL-----GAEMVIN-----TSKYGYK 447

400 VYTEEDLGDLKLLKQLTWEGASQSWYNLWKEFRSYLQPRNPGRELNIRIRVKSGETQK 459

448 STFCSD 454

Db 460 LIFCTED 466

RESULT 12

US-08-985-492-10

; Sequence 10, Application US/08985492

; Patent No. 6395530

; GENERAL INFORMATION:

; APPLICANT: Jaye, Michael C.

; APPLICANT: Doan, Kim-Anh T.

; APPLICANT: Krawiec, John A.

; APPLICANT: Lynch, Kevin J.

; APPLICANT: Amin, Dilip V.

; APPLICANT: South, Victoria J.

; TITLE OF INVENTION: LIG POLYPEPTIDES OF THE TRIACYLGLYCEROL

; TITLE OF INVENTION: LIPASE FAMILY, AND COMPOSITIONS AND METHODS FOR THEIR USE

; TITLE OF INVENTION: IN ENZYMATIC HYDROLYSIS, AND PROTEIN AND GENE THERAPIES

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Rhone-Poulenc Rorer Inc.

; STREET: 500 Arcola Rd. 3C43

; CITY: Collegeville

; STATE: PA

; COUNTRY: USA

; ZIP: 19426

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/985,492

; FILING DATE:

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: Fehner Ph.D., Paul F.

; REGISTRATION NUMBER: 35,135

; REFERENCE/DOCKET NUMBER: A2582-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (610)454-3839

; TELEFAX: (610)454-3808

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 345 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-985-492-10

Query Match 18.3%; Score 462.5; DB 4; Length 345;

Best Local Similarity 34.0%; Pred. No. 5.2e-41;

Matches 123; Conservative 46; Mismatches 134; Indels 59; Gaps 15;

Qy 20 VCYERLCGKDFGKPLWTRFTSTELVGLPWSPE-----KINTRELLYTIH 62

Db 13 LCY-----CFAAGSP-----VFPGEGRLEDKHKPKATQTEVKPSVRNLRATSK 57

Qy 63 NP---NAYQEI SAVNSTIQASYFGTDKITRINAGWKTDG---KWQDMCNVLLQLE-D 115

Db 58 DPEHEGCV--LSVGHSGPLEDCSFNWTAKTFIIGHWTMSGIFENWJHKLVSALHTREKD 115

Qy 116 INCINLDWINGSRE-YIHAVNNLRVVGAEVAYFIDVLMKKFEYSPSKVHLIGHSLGAHLA 174

Db 116 ANVVVDWDLPLAHLQLYTDAVNNTRVVGHSIARMLDWLQEKDQDFSLGNVHLIGYSLGAHVA 175

Qy 175 GEAGSRIPG-LGRITGLDPAGPFPHNTPEKVRLDPSDANFVDVIHTNAARILFELGVGTI 233

Db 176 GYAGNFVKGTGVRITGLDPAGPFEGADIKRLSPDDADFDVLHTYTRS--FGLSIGIQ 233

Qy 234 DACGHLDFYPNGKHPGCG--EDLITPLKFNFNAYKEMASFFDCNHARSYQFYAESIL 291

Db 234 MPVGHIDIYPNGDFQPGGLNDVLGSI-----AY-GTITEVVKCEHERAVHLFVDSL 286

292 NPD-AFIAYPCRSYTSFKAGNCFPCSKGCGPTMGHFAFRPHFKNMKTNGSHYFLNTGSL 350  
 287 NODKPSFAFQCTDSNRFFKKGICLSCKRNCNSIGYNAKVRNK-----RNSKMYLKTRAGM 342  
 351 PF 352  
 343 PF 344

r 13  
 -985-492-4  
 jence 4, Application US/08985492  
 snt No. 6395530

VERAL INFORMATION:  
 APPLICANT: Jaye, Michael C.  
 APPLICANT: Doan, Kim-Anh T.  
 APPLICANT: Krawiec, John A.  
 APPLICANT: Lynch, Kevin J.  
 APPLICANT: Amin, Dilip V.  
 APPLICANT: South, Victoria J.  
 TITLE OF INVENTION: LLG POLYPEPTIDES OF THE TRIACYLGLYCEROL  
 TITLE OF INVENTION: LIPASE FAMILY, AND COMPOSITIONS AND METHODS FOR THEIR USE  
 TITLE OF INVENTION: IN ENZYMATIC HYDROLYSIS, AND PROTEIN AND GENE THERAPIES  
 NUMBER OF SEQUENCES: 31  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Rhone-Poulenc Rorer Inc.  
 STREET: 500 Arcola Rd. 3C43  
 CITY: Collegeville  
 STATE: PA  
 COUNTRY: USA  
 ZIP: 19426  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/985,492  
 FILING DATE:  
 CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fehner Ph.D., Paul F.  
 REGISTRATION NUMBER: 35,135  
 REFERENCE/DOCKET NUMBER: A2582-US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (610)454-3839  
 TELEFAX: (610)454-3808  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 353 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 -985-492-4

Query Match 18.3%; Score 462.5; DB 4; Length 353;  
 Local Similarity 34.0%; Pred. No. 5.4e-41;  
 Matches 123; Conservative 46; Mismatches 134; Indels 59; Gaps 15;

20 VCVERLGCDFKGLPWTRTFSTELVGLPWSPE-----KINTRFLLYTH 62  
 13 LCY-----CFAAGSP-----VPGFEGRELEDKHKPKATQTEVPSVRFNLRISK 57  
 63 NP---NAYQELISAVNSSTIQASYFGTDKLTIRINIAWKTDG---KQWDMCNVLLQLE-D 115  
 58 DPEHEGCV--LSVGHSPLEDCSFNMATKFFIIHGWMTSGIFENWLKLVLSALHTRKD 115  
 116 INCINLDWINGSRE-YHAVNNLVGAEVAYFDVLMKKFYSPSKVHLIGHSGAHUA 174  
 116 ANVVVDVPLPLAHLQLYTDVNNTRVWGHSGIARMLDWLQKQDPSLGNVHLIGYSLGAHYA 175  
 175 GEAGSRIPG-LGRITGLDPAGFPFHNTPKVRLDPSDANVDVHTNAARILFELGVGTI 233

Db 176 VGAGNFVKGTVGRITGLDPAGMPEGADIKRLSPDDADFDVLFHTVTRS--FCLSLGIQ 233  
 QY 234 DACGHLDPYNGGKMGPC--EDLITPLKFNFNAYKEMASPFDCNHARSYQFYAESIL 291  
 Db 234 MPVGHIDIYPNGGDFQPGCGGLNDVLGSI-----AY-GTITEVVKCEHERAVHLFVDSL 286  
 QY 292 NPD-AFIAYPCRSYTSFKAGNCFPCSKGCGPTMGHFAFRPHFKNMKTNGSHYFLNTGSL 350  
 Db 287 NODKPSFAFQCTDSNRFFKKGICLSCKRNCNSIGYNAKVRNK-----RNSKMYLKTRAGM 342  
 QY 351 PF 352  
 Db 343 PF 344

## RESULT 14

US-08-985-492-6  
 Sequence 6, Application US/08985492  
 Patent No. 6395530  
 GENERAL INFORMATION:  
 APPLICANT: Jaye, Michael C.  
 APPLICANT: Doan, Kim-Anh T.  
 APPLICANT: Krawiec, John A.  
 APPLICANT: Lynch, Kevin J.  
 APPLICANT: Amin, Dilip V.  
 APPLICANT: South, Victoria J.  
 TITLE OF INVENTION: LLG POLYPEPTIDES OF THE TRIACYLGLYCEROL  
 TITLE OF INVENTION: LIPASE FAMILY, AND COMPOSITIONS AND METHODS FOR THEIR USE  
 TITLE OF INVENTION: IN ENZYMATIC HYDROLYSIS, AND PROTEIN AND GENE THERAPIES  
 NUMBER OF SEQUENCES: 31  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Rhone-Poulenc Rorer Inc.  
 STREET: 500 Arcola Rd. 3C43  
 CITY: Collegeville  
 STATE: PA  
 COUNTRY: USA  
 ZIP: 19426  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/985,492  
 FILING DATE:  
 CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fehner Ph.D., Paul F.  
 REGISTRATION NUMBER: 35,135  
 REFERENCE/DOCKET NUMBER: A2582-US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (610)454-3839  
 TELEFAX: (610)454-3808  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 354 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-985-492-6

Query Match 18.3%; Score 462.5; DB 4; Length 354;  
 Best Local Similarity 34.0%; Pred. No. 5.4e-41;  
 Matches 123; Conservative 46; Mismatches 134; Indels 59; Gaps 15;  
 QY 20 VCVERLGCDFKGLPWTRTFSTELVGLPWSPE-----KINTRFLLYTH 62  
 Db 13 LCY-----CFAAGSP-----VPGFEGRELEDKHKPKATQTEVPSVRFNLRISK 57  
 QY 63 NP---NAYQELISAVNSSTIQASYFGTDKLTIRINIAWKTDG---KQWDMCNVLLQLE-D 115  
 Db 58 DPEHEGCV--LSVGHSPLEDCSFNMATKFFIIHGWMTSGIFENWLKLVLSALHTRKD 115

116 INCINLDWINGSRE-YIHAVNNLRVGAEVAYFIDVLMKKFEYSPSKVHLIGHSLGAHLA 174  
 116 ANVVVDWLPLAHQLYTDVANNTRVVGHSIARMLDWLQEKDDFSLGNVHLIGYSLGAHVA 175  
 175 GEASRIPLG-LGRITGLDPAGPFFHTPKPEVRLDPSDANFVDVHTNNAARILFELGVGTI 233  
 176 GYAGNFVKGTVGRTITGLDPAGPMFEGADIRKLSPPDADFDVLHTYTRS--FGLSIGIQ 233  
 234 DACGHLDFYNGGKHMFGC--EDLIITPLKFNENAYKKEMASFFDCNHAARSYQVVAESIL 291  
 234 MPVGHIDIYNGGDFQFGGGLNDVLSI-----AY-GTITVVKCEHERAVHLFVDSLIV 286  
 292 NPD-AFTAYPCRSYTSKAGNCFCEGCPWGHFADRFHFNKMKNGSHYFLNTGSL 350  
 287 NQDKPSFAFOCTDSNREFKKGICLSCRNCNSGYNAKKVRNK-----RNSKMYLXTRAGM 342  
 351 PF 352  
 343 PF 344

JT 15  
 8-985-492-13  
 pence 13. Application US/08985492  
 ent No. 6395530  
 INERAL INFORMATION:  
 APPLICANT: Jaye, Michael C.  
 APPLICANT: Doan, Kim-Anh T.  
 APPLICANT: Krawiec, John A.  
 APPLICANT: Lynch, Kevin J.  
 APPLICANT: Amin, Dilip V.  
 APPLICANT: South, Victoria J.  
 TITLE OF INVENTION: LLG POLYPEPTIDES OF THE TRIACYLGLYCEROL  
 TITLE OF INVENTION: LIPASE FAMILY, AND COMPOSITIONS AND METHODS FOR THEIR USE  
 TITLE OF INVENTION: IN ENZYMATIC HYDROLYSIS, AND PROTEIN AND GENE THERAPIES  
 NUMBER OF SEQUENCES: 31  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Rhone-Poulenc Rorer Inc.  
 STREET: 500 Arcola Rd. 3043  
 CITY: Collegeville  
 STATE: PA  
 COUNTRY: USA  
 ZIP: 19426  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/985,492  
 FILING DATE:  
 CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fehner Ph.D., Paul F.  
 REGISTRATION NUMBER: 35,135  
 REFERENCE/DOCKET NUMBER: A2582-US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (610)454-3839  
 TELEFAX: (610)454-3808  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 472 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 8-985-492-13

ry Match 17.4%; Score 440.5; DB 4; Length 472;  
 t Local Similarity 28.1%; Pred. No. 2.1e-38;  
 ches 127; Conservative 79; Mismatches 195; Indels 51; Gaps 16;

QY 34 WTRFTETELVGLPWPBPEK-----INTRFLLYTTHNPAYQE-----ISAVNSSTIOASYF 83  
 Db 14 WQSLTASRGVAAADQDRDFIDIESKALRT---PETAEDTCHLIPGV-AESVATCHF 69  
 QY 84 GTDKITRINIAGWKTDG---KWQDMCNVLQLE-DINCINLDWINGSREYIHAVNNLRV 139  
 Db 70 NHSSKTFMVIHGWTVTGMYESWVVKLVAALKYKREPDSNVIVDWLSRAQEHYVPVSAGYTK 129  
 QY 140 VQAEVAYFDVLMKKFEYSPSKVHLIGHSLGAHLAGEAGSRI-PGLGRITGLDPAGPFFH 198  
 Db 130 VGODVARFINWMEEFNPLDNVHLGYSLGAHAGIAGSLTNKKVNRITGLDPAGNFE 189  
 QY 199 NTPKEVRLDPSDANFVDVHTNNAARILFELG-----VGTIDACGHLDFYPNGGKHMFGCE 253  
 Db 190 YA-EAPRLSPDDADFDVLHT-----FTRGSPGRSIGIQPKPVGHVDIYPNGGTFQPGCN 242  
 QY 254 --DLITPLKFNENAYKKEMASFFDCNHAARSYQVVAESILNPD-AFTAYPCRSYTSFKAG 310  
 Db 243 IGEAIRVIAERGLGDVDQ-----LKCSEHSIHLFIDSLLEENPSPKAYRCSSEAFK 297  
 QY 311 NCFPCSKGECPTMGHFADRFHFNKMKNGSHYFLNTGSLSPFARWRHKLKSLGSEV-- 368  
 Db 298 LCLSCRNKRNCLNGLYEINKVRK-----RSSKMYLKTRSQMPYKVFHYQVKIHFSGTSET 353  
 QY 369 -TQGTVFLVGAIGKTGEFALVSGKLEPFGMTYTKLIDADVNGNITSVQFIWKHLFED 427  
 Db 354 HTNQAFELISLYGTVAESENIPFTLPEVSTNKTYSFLIYTYEVDIGELMLKLKWKSDSYFS 413  
 QY 428 -----SQNKLGAEMVINTSGKYGYKSTFCSDQ 454  
 Db 414 WSDMWSPPGFAIQKIRVXKAGETQKKVIFCSRE 445

Search completed: January 6, 2004, 11:24:08  
 Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

rotein - protein search, using sw model

n: January 6, 2004, 08:50:18 ; Search time 51 Seconds  
(without alignments)  
1453.437 Million cell updates/sec

US-10-038-517-2

ct score: 2527

nce: 1 MGWIVAFLEFGTSGRKEV.....STFCSQDMGNILQNKPC 467

ng table: BLOSUM62

Gapop 10.0 , Gapext 0.5

hed: 1107863 seqs, 158726573 residues

number of hits satisfying chosen parameters: 1107863

um DB seq length: 0

um DB seq length: 2000000000

processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

ase : A\_Geneseq\_19Jun03:\*

- 1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Score	Query Match	Length	ID	Description
1 2527	100.0	467	23 ABB77987	Amino acid sequenc
2 2527	100.0	467	23 ABB79039	Human lipase prote
3 2523	99.8	467	22 AAE00508	Human lipase prote
4 2276	90.1	461	23 AAE17302	Human pancreatic l
5 1937	76.7	368	23 ABB08004	Human lipid metabo
6 1878.5	74.3	462	23 ABB53289	Human polypeptide
7 1456	57.6	276	23 ABB97589	Novel human protei
8 1191	47.1	469	23 ABB79040	Mycocastor corpus p
9 1180	46.7	467	23 ABB79041	Mouse pancreatic i

10	1172	46.4	467	23	ABB79042	Rat pancreatic lip
11	1162	46.0	467	23	ABB06317	Plrpl amino acid s
12	1159	45.9	470	14	AA30741	GPI(+) mutant with
13	1156	45.7	452	14	AA30737	Guinea pig pancrea
14	1152	45.6	469	22	ABG16896	Novel human diagno
15	1152	45.6	469	22	ABG06318	Plrpl2 amino acid s
16	1152	45.6	477	21	AA34287	Human pancreatic c
17	1143.5	45.3	474	22	ABG16895	Novel human diagno
18	1115	44.1	465	23	ABB06321	Human lipase amino
19	1115	44.1	465	23	AAU87690	Human pancreatic t
20	1115	44.1	473	21	AA34115	Human pancreatic c
21	1091.5	43.2	450	14	AA30738	Human pancreatic l
22	1088.5	43.1	432	14	AA34294	HPL(-) mutant with
23	980	38.8	392	21	AA300075	Human secreted pro
24	899.5	35.6	336	14	AA30740	C-terminally truca
25	886.5	35.1	470	22	ABG16892	Novel human diagno
26	886	35.1	319	14	AA30739	C-terminally trunc
27	879.5	34.8	336	14	AA34293	C-terminally trunc
28	866	34.3	319	14	AA34292	C-terminally trunc
29	555.5	22.0	952	22	ABB62733	Drosophila melanog
30	534	21.1	498	23	ABB06319	Lipase amino acid
31	534	21.1	499	23	AAO15887	Human hepatic lipa
32	525	20.8	500	20	AAV23760	Mouse endothelial
33	515.5	20.4	452	23	AAE17316	Human lipase prote
34	515.5	20.4	500	19	AAW59792	Amino acid sequenc
35	515.5	20.4	500	20	AAV23759	Human endothelial
36	515.5	20.4	500	21	AA19178	Human LIPG, a tria
37	515.5	20.4	500	23	AAO14635	Human lipase endot
38	515.5	20.4	500	24	ABU03569	Angiogenesis-assoc
39	513	20.3	451	22	AAE10996	Human lipid metabo
40	513	20.3	451	22	AAE10996	Human lipid metabo
41	513	20.3	451	23	ABB73560	Human lipase 18892
42	513	20.3	451	23	ABB06316	Human phospholipas
43	500	19.8	540	22	ABB08952	Human novel lipase
44	462.5	18.3	345	19	AAW59793	Drosophila melanog
45	462.5	18.3	345	21	AA19179	Amino acid sequenc
						The catalytic doma

ALIGNMENTS

RESULT 1

ABB77987  
ID ABB77987 standard; Protein, 467 AA.

XX ABB77987;

DT 22-OCT-2002 (first entry)

DE Amino acid sequence of a human lipase,

Human; lipase; enzyme; mental disorder; disease; gene.

OS Homo sapiens.

PN WO200253753-A2.

PD 11-JUL-2002.

PF 03-JAN-2002; 2002WO-US00223.

PR 05-JAN-2001; 2001US-259830P.

XX (LEXI-) LEXICON GENETICS INC.

PI Yu X, Turner CA;

XX WPI; 2002-583623/62.

DR N-PSDB; ABL59436.

XX Nucleic acid molecule encoding a human lipase, useful in therapeutic,  
PT diagnostic and pharmacogenomic applications, as DNA markers for  
PT restriction fragment length polymorphism analysis and in forensic

biology -

Claim 4; Page 35-36; 36pp; English.

The present sequence represents a human lipase polypeptide. The lipase polynucleotides and polypeptides are useful in therapeutic, diagnostic and pharmacogenomic applications, and for identifying compounds that modulate, i.e., act as agonists or antagonists of the gene expression or gene product activity. They are also useful in the human lipase protein gene regulation and/or as antisense primers in amplification reactions of the nucleic acid sequences, for detecting mutant human proteins or inappropriately expressed proteins for the diagnosis of disease, for screening for drugs effective in treating the symptomatic or phenotypic manifestations of perturbing the normal function of the protein in the body, for generating transgenic animals and antibodies, for identifying other cellular gene products related to the protein, and as reagents in assays for screening for compounds that can be used as pharmaceutical agents in the therapeutic treatment of mental, biological or medical disorders and diseases.

Sequence 467 AA;

Query Match 100.0%; Score 2527; DB 23; Length 467;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-244;  
 Mismatches 0; Indels 0; Gaps 0;

1 MLGIWIVAFPFPGTSRGKVCYERLGCDFDGLPWTFTSTELVGLPWSPEKINTRFLLYT 60  
 1 MLGIWIVAFPFPGTSRGKVCYERLGCDFDGLPWTFTSTELVGLPWSPEKINTRFLLYT 60  
 61 IHNPNAYQEISAVNSSTIQASYFGTDKITRINIAGWKTDGKQWDMCNVLLQLEDINCIN 120  
 61 IHNPNAYQEISAVNSSTIQASYFGTDKITRINIAGWKTDGKQWDMCNVLLQLEDINCIN 120  
 121 LDWINGSREYIHAVNNLRVVGAEVAYFIDVLMKKFYSKVKHLIGHSLGAHLAGEAGSR 180  
 121 LDWINGSREYIHAVNNLRVVGAEVAYFIDVLMKKFYSKVKHLIGHSLGAHLAGEAGSR 180  
 181 IPGLGRITGLDPAGPFFHNTPEKVRLDPSDANFVDVHTNAARILFELGVGTIDACGHL 240  
 181 IPGLGRITGLDPAGPFFHNTPEKVRLDPSDANFVDVHTNAARILFELGVGTIDACGHL 240  
 241 FYPNGGKHPGCCDLITPLLKFNFNAYKEMASFFDCNHARSYQFYAESILNPDATAYP 300  
 241 FYPNGGKHPGCCDLITPLLKFNFNAYKEMASFFDCNHARSYQFYAESILNPDATAYP 300  
 301 CRSYTSFKAGNCFCKECCPTMGHFDADRFHFNKMTNGSHYFLNTGSLSPFARWRHKL 360  
 301 CRSYTSFKAGNCFCKECCPTMGHFDADRFHFNKMTNGSHYFLNTGSLSPFARWRHKL 360  
 361 VKLSGSEVTQGTVFLRVGGAIGKTGEFAIVSGKLBPGMTYTKLIDADVNVGNITSVQFIW 420  
 361 VKLSGSEVTQGTVFLRVGGAIGKTGEFAIVSGKLBPGMTYTKLIDADVNVGNITSVQFIW 420  
 421 KHLFEDSQNKLGAEMWINTSGKYGYKSTFCSQDIMGPNILQNLKPC 467  
 421 KHLFEDSQNKLGAEMWINTSGKYGYKSTFCSQDIMGPNILQNLKPC 467

T 2  
 039

AB579039 standard; Protein; 467 AA.

AB579039;

01-AUG-2002 (first entry)

Human lipase protein sequence SEQ ID NO:2.

Human; lipase; enzyme; therapeutic.

Homo sapiens.

PN US2002052034-A1.  
 XX  
 PD 02-MAY-2002.  
 XX  
 PF 14-DEC-2000; 2000US-0735933.  
 XX  
 PR 28-SEP-2000; 2000US-235925P.  
 XX  
 PA (GUEG/) GUEGLER K.  
 PA (WEBB/) WEBSTER M.  
 PA (KETCH/) KETCHUM K A.  
 PA (DFRA/) DI FRANCESCO V.  
 PA (SEAS/) BEASLEY E M.  
 XX  
 PI Guegler K, Webster M, Ketchum KA, Di Francesco V, Beasley EM;  
 DR WPI; 2002-443697/47.  
 DR N-PSDB; ABN97363, ABN97364.  
 XX  
 PT A lipase protein identified from the human genome is related to the  
 PT pancreatic lipase subfamily, and is useful to provide new drug targets  
 PT in the discovery of therapeutics for lipase-mediated disease  
 XX  
 PS Claim 1; Fig 2; 85pp; English.  
 XX  
 CC The present sequence represents a human lipase protein. Molecules from  
 CC the present invention can be used as models for the development of human  
 CC therapeutic targets, aid in the identification of therapeutic proteins  
 CC and serve as targets for the development of human therapeutic agents  
 CC that modulate lipase activity in cells and tissues that express lipase.  
 XX  
 SQ Sequence 467 AA;

Query Match 100.0%; Score 2527; DB 23; Length 467;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-244;  
 Mismatches 0; Indels 0; Gaps 0;  
 Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLGIWIVAFPFPGTSRGKVCYERLGCDFDGLPWTFTSTELVGLPWSPEKINTRFLLYT 60  
 Db 1 MLGIWIVAFPFPGTSRGKVCYERLGCDFDGLPWTFTSTELVGLPWSPEKINTRFLLYT 60  
 QY 61 IHNPNAYQEISAVNSSTIQASYFGTDKITRINIAGWKTDGKQWDMCNVLLQLEDINCIN 120  
 Db 61 IHNPNAYQEISAVNSSTIQASYFGTDKITRINIAGWKTDGKQWDMCNVLLQLEDINCIN 120  
 QY 121 LDWINGSREYIHAVNNLRVVGAEVAYFIDVLMKKFYSKVKHLIGHSLGAHLAGEAGSR 180  
 Db 121 LDWINGSREYIHAVNNLRVVGAEVAYFIDVLMKKFYSKVKHLIGHSLGAHLAGEAGSR 180  
 QY 181 IPGLGRITGLDPAGPFFHNTPEKVRLDPSDANFVDVHTNAARILFELGVGTIDACGHL 240  
 Db 181 IPGLGRITGLDPAGPFFHNTPEKVRLDPSDANFVDVHTNAARILFELGVGTIDACGHL 240  
 QY 241 FYPNGGKHPGCCDLITPLLKFNFNAYKEMASFFDCNHARSYQFYAESILNPDATAYP 300  
 Db 241 FYPNGGKHPGCCDLITPLLKFNFNAYKEMASFFDCNHARSYQFYAESILNPDATAYP 300  
 QY 301 CRSYTSFKAGNCFCKECCPTMGHFDADRFHFNKMTNGSHYFLNTGSLSPFARWRHKL 360  
 Db 301 CRSYTSFKAGNCFCKECCPTMGHFDADRFHFNKMTNGSHYFLNTGSLSPFARWRHKL 360  
 QY 361 VKLSGSEVTQGTVFLRVGGAIGKTGEFAIVSGKLBPGMTYTKLIDADVNVGNITSVQFIW 420  
 Db 361 VKLSGSEVTQGTVFLRVGGAIGKTGEFAIVSGKLBPGMTYTKLIDADVNVGNITSVQFIW 420  
 QY 421 KHLFEDSQNKLGAEMWINTSGKYGYKSTFCSQDIMGPNILQNLKPC 467  
 Db 421 KHLFEDSQNKLGAEMWINTSGKYGYKSTFCSQDIMGPNILQNLKPC 467

RESULT 3  
 AAB00508  
 ID AAB00508 standard; Protein; 467 AA.

AAE00508;

19-JUN-2001 (first entry)

Human lipase protein, MLip-1.

Human, lipase; MLip-1; antiinflammatory; antilipemic; cardiatic; therapy; anorectic; cholesterol biosynthesis; lipid metabolism; diabetes; obesity; pancreatic disorder; nutritional disorder; malabsorption; malnutrition; metabolic disorder; hyperlipidaemia; hypolipidaemia; abdominal lesion; lipidoses; pancreatic duct obstruction; Gaucher's disease; peritonitis; atherosclerosis; Niemann-Pick disease; arteriosclerosis; pancreatitis; coronary artery disease; perforated peptic ulcer; intestinal obstruction; linoleic acid deficiency.

Homo sapiens.

Key Location/Qualifiers

Peptide 1..17

Modified-site 13..18

Protein 18..467

Modified-site 31..36

Domain 42..343

Modified-site 125..128

Modified-site 141..146

Active-site 162..171

Modified-site 170..175

Modified-site 189..194

Modified-site 231..236

Modified-site 338..341

Domain 355..467

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

Modified-site 439..442

PA (MILL-) MILLENIUM PHARM INC.

XX Khodadoust M, Kapeller-libermann R;

XX WPI; 2001-266302/27.

XX N-PSDB; AAD03848.

XX Novel isolated lipase, MLip-1, useful for identifying compounds which

XX modulate MLip-1 activity, for treating conditions or disorders

XX PT associated with aberrant activity of MLip-1 such as hypolipidemia,

XX obesity

XX Claim 9; Fig 1; 112pp; English.

XX The present sequence is human lipase protein, MLip-1. MLip-1

XX catalyse the formation and cleavage of ester bonds between fatty acyl

XX moieties and glyceride moieties. It is involved in dietary fat

XX degradation and absorption, cholesterol biosynthesis and maintenance

XX of plasma lipid and lipoprotein levels. MLip-1 is also useful for

XX identifying compounds which modulates its activity and expression. These

XX compounds are useful for preventing, diagnosing or treating disorders

XX relating to inappropriate lipid metabolism and aberrant pancreatic

XX function such as diabetes, obesity, nutritional disorders (e.g. lipid

XX malabsorption and malnutrition), metabolic disorders (particularly

XX including lipid metabolism anomalies such as hyperlipidaemia of types I

XX to V and hypolipidaemia), pancreatitis, obstruction of the pancreatic

XX duct, various lipidoses (e.g. Gaucher's disease and Niemann-Pick

XX disease), atherosclerosis, arteriosclerosis, coronary artery disease,

XX perforated peptic ulcer, abdominal lesions, intestinal obstruction,

XX peritonitis and linoleic acid deficiency. MLip-1 and its modulators can

XX be used in screening assays, detection assays (e.g. chromosomal mapping,

XX tissue typing and forensic biology assays), predictive medicine (e.g.

XX diagnostic assays, prognostic assays, monitoring of clinical trials and

XX pharmacogenomic applications) and methods of treatment (e.g. therapeutic

XX and prophylactic methods).

XX Sequence 467 AA;

XX

SQ

Query Match 99.8%; Score 2523; DB 22; Length 467;

Best Local Similarity 99.8%; Pred. No. 9.6e-244;

Matches 466; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 MGIIWIVAFVFFGTSRGKVCYERLGCFCGDLGFWTFTSTELVGLPWSPEKINTRELLYT 60

1 MGIIWIVAFVFFGTSRGKVCYERLGCFCGDLGFWTFTSTELVGLPWSPEKINTRELLYT 60

61 IHNPNAYQESAVNSSTIQASYFGTDKTRINIAAGWKTGKQWDCNVLLQLEDINCIN 120

61 IHNPNAYQESAVNSSTIQASYFGTDKTRINIAAGWKTGKQWDCNVLLQLEDINCIN 120

121 LDWINGSREVIHAVNNLRVVGAEVAFIDVLMKKFYSKVKHLIGHSLGAHLAGEAGSR 180

121 LDWINGSREVIHAVNNLRVVGAEVAFIDVLMKKFYSKVKHLIGHSLGAHLAGEAGSR 180

181 IPGLGRITGLDPAGPFFHNTPEKVRDPSDANFVDVHTNAARILPELGVGTIDACGHLD 240

181 IPGLGRITGLDPAGPFFHNTPEKVRDPSDANFVDVHTNAARILPELGVGTIDACGHLD 240

241 FYPNGGKMPGCCEDLITPLKFNFAKXEMASFFDCNHRASYQFVAESILNPDAFIAP 300

241 FYPNGGKMPGCCEDLITPLKFNFAKXEMASFFDCNHRASYQFVAESILNPDAFIAP 300

301 CRSYTSFKAGNCFCSKEGCTMGHFAFRHFKNMKTNGSHYFLNTGSLSPFARWRHKL 360

301 CRSYTSFKAGNCFCSKEGCTMGHFAFRHFKNMKTNGSHYFLNTGSLSPFARWRHKL 360

361 VKLSGSEVTGCTVFLVGGAGIKGTGFAIVSGKLEPGMTYTKLIDADVNVGNITSVQFIW 420

361 VKLSGSEVTGCTVFLVGGAGIKGTGFAIVSGKLEPGMTYTKLIDADVNVGNITSVQFIW 420

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467

421 KGHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDIMGNILQNLKPC 467



JT 4  
7302

AAE17302 standard; Protein; 461 AA.

AAE17302;

18-APR-2002 (first entry)

Human pancreatic lipase protein, sbg237163LIPASE.

Human; therapy; wound healing disorder; vaccine; cancer; infection; autoimmune disorder; haematopoietic disorder; inflammation; arthritis; Parkinson's disease; Huntington's chorea; schizophrenia; antiarrhythmic; multiple sclerosis; Alzheimer's disease; analgesic; cardiant; asthma; ischaemia; stroke; AIDS; bone disease; atherosclerosis; brain disorder; depression; cardiovascular disease; myocardial infarction; renal failure; respiratory disease; liver disorder; Fanconi's syndrome; spleen disorder; type II diabetes mellitus; skeletal muscle disorder; immunosuppressive; hyperplasia; renal disease; hypoglycaemia; gastrointestinal disease; neutropenic; cirrhosis; Hodgkin's disease; neuroleptic; antiinflammatory; haemostatic; vulvar; anticonvulsant; antirheumatic; neuroprotective; nephrotropic; hypotensive; vasotropic; cytostatic; cerebroprotective; allergy; pancreatic lipase.

Homo sapiens.

WO200198342-A1.

27-DEC-2001.

22-JUN-2001; 2001WO-US19929.

22-JUN-2000; 2000US-213156P.

22-JUN-2000; 2000US-213156P.

(SMIK) SMITHKLINE BEECHAM CORP.

(SMIK) SMITHKLINE BEECHAM PLC.

(GLAX) GLAXO GROUP LTD.

Agarwal P, Cogswell JP, Kabnic KS, Lai Y, Martensen SA; Murdock PR, Smith RF, Strum JC, Xiang Z, Xie Q, Rizni SK;

WPI; 2002-139783/18.

N-PSDB; AAD2777.

Novel secreted and membrane-associated polypeptides and polynucleotides useful for preventing, ameliorating or correcting dysfunction or disease including diabetes, cancer, hypertension and growth abnormalities

Claim 1; Page 99-100; 138pp; English.

The invention relates to secreted and membrane-associated polypeptides and polynucleotides. The sequences of the invention are useful in diagnostic assays for detecting diseases associated with inappropriate activity or levels of these polynucleotides, and in identifying their agonists and antagonists that are potentially useful in therapy. The sequences of the invention are useful as vaccines for inducing immunological response. The sequences of the invention are useful for treating cancers, infections, autoimmune disorders, haematopoietic disorders, wound healing disorders, cholesterol ester storage disease, inflammation, congenital muscular dystrophy, junctional epidermolysis bullosa, Parkinson's disease, Huntington's chorea, multiple sclerosis, viral and bacterial infections, Alzheimer's disease, asthma, arthritis, allergies, schizophrenia, sbg44245PROs-associated disorders, septicemia, psoriasis, inflammatory bowel disease, transplant rejection, graft versus host disease, ischaemia, stroke, acute respiratory disease syndrome, restenosis, brain injury, AIDS, bone diseases, atherosclerosis, brain disorders including paraspranuclear palsy, myotonic dystrophy, depression, anxiety disorders and sleep disorders, cardiovascular diseases including congestive heart failure and myocardial infarction,

CC respiratory diseases including chronic obstructive pulmonary disease, CC acute bronchitis and adult respiratory distress syndrome, liver disorders CC including hypercholesterolaemia, hypertriglyceridaemia, cirrhosis, viral CC and non-viral hepatitis, type II diabetes mellitus, renal disease CC including acute and chronic renal failure, glomerulonephritis, Fanconi's CC syndrome, cystinuria, skeletal muscle disorders including hypoglycaemia CC and tendinitis, gastrointestinal diseases including intestinal CC obstruction and tropical sprue, spleen disorders including hypersplenism, CC Hodgkin's disease and malignant lymphoma, testicular cancer, male CC reproductive diseases including low testosterone and male infertility. CC The present sequence is human pancreatic lipase.

XX SQ Sequence 461 AA;

Query Match 90.1%; Score 2276; DB 23; Length 461;  
Best Local Similarity 91.5%; Pred. No. 5.5e-219;  
Matches 431; Conservative 8; Mismatches 18; Indels 14; Gaps 3;

QY 1 M LGIWIIVAFPGTGRGKVCYERLCGFCGDPWTRTSTELVGLPMSPEKINTRFLIYT 60

Db 1 M LGIWIIVAFPGTGRGKVCYERLCGFCGDPWTRTSTELVGLPMSPEKINTRFLIYT 60

QY 61 IHNPNAYQISAVNSTTQASVFGTDKITRINAGWKTDGKQDMCNVLLQLEDINCIN 120

Db 61 IHNPNAYQISAVNSTTQASVFGTDKITRINAGWKTDGKQDMCNVLLQLEDINCIN 120

QY 121 LWMNGSREYIHAVNNLRVVGAEVAFIDVLMKKPEYSPKVLHIGSLGAHLAGEAGSR 180

Db 121 LWMNGSREYIHAVNNLRVVGAEVAFIDVLMKKPEYSPKVLHIGSLGAHLAGEAGSR 180

QY 181 IPGLGRITGLDPAGPFFHNTPKPEVRLDPSDFVVDVHTNAARILFELGVGTIDACGHL 240

Db 181 IPGLGRITGLDPAGPFFHNTPKPEVRLDPSDFVVDVHTNAARILFELGVGTIDACGHL 240

QY 241 FYPNGGKHPGCCDLITPLKFNFNAYKKEVAFIDVLMKKPEYSPKVLHIGSLGAHLAGEAGSR 300

Db 241 FYPNGGKHPGCCDLITPLKFNFNAYKKEVAFIDVLMKKPEYSPKVLHIGSLGAHLAGEAGSR 300

QY 301 CRSYTSFKAGNCFKSCGCTPMGHFADRFHFKMKNKTSVHVLNTG 356

Db 301 CRSYTSFKAGNCFKSCGCTPMGHFADRFHFKMKNKTSVHVLNTG 356

QY 357 HKLSVKLSGSEVTQGTFLRVGGAIGKTGEFAIVSGKLEPGMTYTKLIDADVNVGNITSV 416

Db 357 HKLSVKLSGSEVTQGTFLRVGGAIGKTGEFAIVSGKLEPGMTYTKLIDADVNVGNITSV 416

QY 417 QFIWKHLFEDSQNKLGAEMVINTSGKYKSTFCSDIMGNILQNLKPC 467

Db 417 QFIWKHLFEDSQNKLGAEMVINTSGKYKSTFCSDIMGNILQNLKPC 467

RESULT 5

ABB08004

ID ABB08004 standard; Protein; 368 AA.

XX ABB08004;

XX 27-AUG-2002 (first entry)

DE Human lipid metabolism enzyme (LME)-4 (Id: 7482937CD1).

XX Human; lipid metabolism enzyme; LME; cytostatic; neuroprotective; XX nototropic; cerebroprotective; antiparkinsonian; antialzheimers; vaccine; XX antileukemic; antimicrobial; anti-AIDS; cardiovascular; antianginal; XX gene therapy; protein therapy; enzyme.

OS Homo sapiens.

XX WO200229036-A2.

XX 11-APR-2002.

PF 05-OCT-2001; 2001WO-US31302.

06-OCT-2000; 2000US-238368P.  
13-OCT-2000; 2000US-240616P.  
02-NOV-2000; 2000US-245719P.  
08-NOV-2000; 2000US-247503P.  
17-NOV-2000; 2000US-249503P.

(INCY-) INCVTE GENOMICS INC.

Harland L, Arvizu C, Das D, Griffin JA, Baughn MR, Ding L,  
Wallia NK, Yao MG, Lu Y, Elliott VS, Thangavelu K, Ramkumar J,  
Lal PG, Tribouley CW;

WPI; 2002-315862/35.  
N-PSDB; ABJ60540.

Lipid Metabolism Enzymes and nucleic acids, useful for preventing,  
diagnosing and treating e.g. cancer, Alzheimer's disease and  
Creutzfeld-Jakob disease

Claim 1; Page 115-116; 127pp; English.

The invention relates to human lipid metabolism enzymes (LMEs) and  
encoding polynucleotides. The LMEs can be expressed by standard  
recombinant technology. The LME polypeptides, polynucleotides and  
modulators may be used in the prevention, diagnosis and treatment of  
diseases associated with inappropriate LME expression such as cancer  
(e.g. myeloma, sarcoma and breast cancer), neurological disorders (e.g.  
Parkinson's, Alzheimer's and multiple sclerosis), microbial infections  
(e.g. Creutzfeld-Jakob disease and Acquired Immune deficiency syndrome  
(AIDS)) and/or cardiovascular disorders (e.g. cardiomyopathy, angina  
pectoris and mitral valve prolapse). The present sequence represents  
the human LME-4 polypeptide.

Sequence 368 AA;

Y Match 76.7%; Score 1937; DB 23; Length 368;  
: Local Similarity 99.2%; Pred. No. 3.8e-185;  
:hes 359; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

106 MCNVLLQLEDINCINLDWNGSRVTHAVNNLRVVGAIVFYDVLNKKFEYSPSKVHLI 165

7 LCKVLQLEDINCINLDWNGSRVTHAVNNLRVVGAIVFYDVLNKKFEYSPSKVHLI 66

166 GHSGLAGHLAGAGSRIPGLGRITGLDPAGPFPHNTPKVEVLDPDSANFVDVHTNAARIL 225

67 GHSGLAGHLAGAGSRIPGLGRITGLDPAGPFPHNTPKVEVLDPDSANFVDVHTNAARIL 126

226 FELGVGTIDACGHLDYPNGSKHMPGCCDLITPLLKNFNAYKKEVASFPDCNARSYQF 285

127 FELGVGTIDACGHLDYPNGSKHMPGCCDLITPLLKNFNAYKKEVASFPDCNARSYQF 186

286 VAESILNPDATFAYPCRSYTSFKAGNCFKSCGCPMWHGFADRFKQKMGSHYFLN 345

187 VAESILNPDATFAYPCRSYTSFKAGNCFKSCGCPMWHGFADRFKQKMGSHYFLN 246

346 TGSLSFPARWARKLSVKLSGSEVTQGTVFLRVGAIGKTGEFAIVSGKLEPGMTYTKLID 405

247 TGSLSFPARWARKLSVKLSGSEVTQGTVFLRVGAIGKTGEFAIVSGKLEPGMTYTKLID 306

406 ADVNVGNITTSVQFIWKHKLFPDSQNKLAGWVINTSGYKYSKTFCSQDIMGPNILQNLK 465

307 ADVNVGNITTSVQFIWKHKLFPDSQNKLAGWVINTSGYKYSKTFCSQDIMGPNILQNLK 366

466 FC 467

367 FC 368

. 6  
:89  
:B53289 standard; Protein; 462 AA.

AC ABB53289;  
XX 12-FEB-2002 (first entry)  
DT XX Human polypeptide #29.  
DE XX  
XX Human; neurotropic; neuroprotective; anticonvulsant; antidepressant;  
KW neuroleptic; tranquiliser; antiarrhythmic; cardiant; antiasthmatic;  
KW antiinflammatory; antilipemic; hepatotropic; virucide; antidiabetic;  
KW nephrotropic; anorectic; cytostatic; vaccine; neurological disease;  
KW cardiovascular disease; respiratory disease; liver disease;  
KW renal disease; skeletal muscle disease; gastrointestinal disease;  
KW placental disease; testicular cancer; male fertility; pancreatic disease.  
XX OS  
XX Homo sapiens.  
XX WO200181363-A1.  
XX 01-NOV-2001.  
XX 26-APR-2001; 2001WO-US13360.  
XX 27-APR-2000; 2000US-199963P.  
PR 11-MAY-2000; 2000US-203336P.  
PR 25-MAY-2000; 2000US-207087P.  
PR 26-MAY-2000; 2000US-207546P.  
XX (SMIX ) SMITHKLINE BEECHAM CORP.  
PA (SMIX ) SMITHKLINE BEECHAM PLC.  
XX Agarwal P, Murdock PR, Rizvi SK, Smith RF, Xiang Z, Kabnick KS;  
PI Lai Y, Xie Q;  
XX WPI; 2002-041392/05.  
DR N-PSDB; ABA90354.  
XX Novel polypeptides and polynucleotides useful as a vaccine for  
PT preventing and treating diseases associated with the polypeptide, e.g.  
PT Alzheimer's disease, dyslipidemia, obesity, diabetes, infertility,  
PT asthma, amnesia  
XX Claim 1; Page 100-101; 116pp; English.  
PS The invention relates to an isolated polypeptide comprising a 277, 480,  
CC 583, 581, 628, 424, 638, 229, 310, 841, 241, 369, 382, 185, 586, 1026,  
CC 844, 782, 262, 394, 471, 485, 286, 533, 495, 350, 619, 490, 462, 255,  
CC 784, 252, 593, 472, 607, 781, 640, 686 or 154 amino acid sequence as  
CC given in the specification. The polypeptides, modulators of the  
CC polypeptides and antibodies against the polypeptides are useful for  
CC treating diseases such as neurological and psychiatric diseases  
CC including Alzheimer's, paraneuronal palsy, Huntington's disease,  
CC myotonic dystrophy, anorexia and depression; cardiovascular diseases  
CC including congestive heart failure, Hodgkin's disease and myocardial  
CC infarction; respiratory diseases including asthma, chronic obstructive  
CC pulmonary disease, cystic fibrosis and adult respiratory distress  
CC syndrome; liver diseases including hypercholesterolaemia, cirrhosis,  
CC viral and nonviral hepatitis, Type II diabetes mellitus, and impaired  
CC glucose tolerance; renal disease including renal failure, acute tubular  
CC necrosis and glomerulonephritis; skeletal muscle diseases including  
CC Eulenburg's disease, hypoglycaemia and obesity; gastrointestinal  
CC diseases including myotonia congenita and intestinal obstruction; lymph  
CC diseases including lymphagietasia; diseases of placenta including  
CC chorionicarcoma; diseases of testes including testicular cancer,  
CC male reproductive diseases including low testosterone and male  
CC infertility; and disease of pancreas including diabetetic ketoacidosis,  
CC Type 1 and 2 diabetes and obesity. The present sequence is a  
CC polypeptide of the invention.  
XX Sequence 462 AA;

Query Match 74.3%; Score 1878.5; DB 23; Length 462;  
Best Local Similarity 76.7%; Pred. No. 4e-179;  
Matches 368; Conservative 15; Mismatches 66; Indels 31; Gaps 3;

```

1  MLGIWIVAFLEFGTSRGEKVCYVERLGCFFKDGJLFWTRTFSELVGLFWSPEKINTRFLLYT 60
1  MLPLMTLS-LLLCAGAVGEKVCYERLGCFFSDGSPWSEITERPLHILPWSPKDVNTRFLLYT 59
61 IHNPNAYQEISAVNSSTIOASVEGTDKITRIINAGWTKDGKQWDMCNVLQLEIDINCIN 120
60 NENPNNFQEISAVNSSTIOASVEGTDKITRIINAGWTKDGKQWDMCNVLQLEIDINCIN 119
121 LDWINGSREYIHAVNNLRVVGAEVAFIDVLMKKKEYSPSKVHLIGHSLGAHAGEAGSR 180
120 LDWINGSREYIHAVNNLRVVGAEVAFIDVLMKKKEYSPSKVHLIGHSLGAHAGEAGSR 179
181 IYGLGRITGLDPAGPPFHNTPKEVRLDPDSANFVDVIHNAARILIFELGVGTIDAGHLUD 240
180 IYGLGRITGKHALQGLEBCATEGYLLSATLANNVAFDYN-----HMD 222
241 FYPNGGKHMPGCEDLITPLL-----KFNFNAYKXEMASFFDCNHARSQVFA 287
223 ATPILPQWMRGTSNTSNPLPVTSSLCMLADJGSVSLVCLWPEMASFFDCNHARSQVFA 282
288 ESILNPDAEIAVPCRSYTSFKAGNCFCSKEGCPWNHGFADRFHPFNKMTNGSHYFLNTG 347
283 ESILNPDAEIAVPCRSYTSFKAGNCFCSKEGCPWNHGFADRFHPFNKMTNGSHYFLNTG 342
348 SLSPFARWEHKL SVKLSGSEVTQGTVFLRVGGAIGKTGBFAIVSGKLBPGMTYTKLIDAD 407
343 SLSPFARWEHKL SVKLSGSEVTQGTVFLRVGGAIVKTGBFAIVSGKLBPGMTYTKLIDAD 402
408 VNVGNITSVQFTWKXHLFSDSONKLGAEMVINTSGKYGKYSTFCSQDINGPNIQLXPC 467
403 VNVGNITSVQFTWKXHLFSDSONKLGAEMVINTSGKYGKYSTFCSQDINGPNIQLXPC 462

```

T 7  
1589

ABB97589 standard; Protein: 276 AA.

ABB97589;

27-JUN-2002 (first entry)

Novel human protein SEQ ID NO: 857.

Human; antianaemic; vulnery; antiinflammatory; immunomodulator; antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy; neuroprotective; antiparkinsonian; protein therapy; SSR; expressed sequence tag.

**Homo sapiens.**

WO200222660-A2.

21-MAR-2002.

10-SEP-2001: 2001WC-IIS26015

11-SEP-2000: 2000US-0659671

(HYSE-) HVSEB TNC

Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
Xue AJ, Yang Y, Wehrman T, Drmanac RT;

WPI: 2002-292408/33.

An isolated polynucleotide for treating diseases associated with its encoded polypeptide such as cancer and multiple sclerosis - N-FSDB; ABN32113.

Claim 20: SEO ID NO 857: 509pp: English.

The present invention provides the protein and coding sequences of 444

CC	novel human proteins. These were isolated from expressed sequences tags
CC	(ESTs). They can be used to stimulate cell growth, to regulate
CC	haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC	e.g. in burn treatment, to regulate the immune system e.g. to treat
CC	multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC	infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC	stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC	e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC	Parkinson's disease. The present sequence is a protein of the invention.
XX	Sequence 276 AA;
XX	

Query Match	57.6%;	Score 1456;	DB 23;	Length 276;
Best Local Similarity	100.0%;	Pred. No. 4.1e-137;		
Matches 269;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	MLGIWVAFLEFCHTSRGEKVEYERLGCFKDGLDPMWTFSTELUGLPSWSEKINTRPLLYT	60
Db	1	MLGIWVAFLEFCHTSRGEKVEYERLGCFKDGLDPMWTFSTELUGLPSWSEKINTRPLLYT	60
Qy	61	IHNPNAYOEISAVNSSTIQASFGTDKITRINIAAGWTKDGKWORDCNVLLQLEDINCIN	120
Db	61	IHNPNAYOEISAVNSSTIQASFGTDKITRINIAAGWTKDGKWORDCNVLLQLEDINCIN	120
Qy	121	LDWINGSREYTHAYNNLRVVGAEVAYFIDVLMKKFYPSPSKVHLIGHSLGAHLAGEAGSR	180
Db	121	LDWINGSREYTHAYNNLRVVGAEVAYFIDVLMKKFYPSPSKVHLIGHSLGAHLAGEAGSR	180
Qy	181	IPGLGRITGLDPAGCFPHNTPEKVRDPDSANFVDVIHTNAARIIFELGVGTIDACGHLD	240
Db	181	IPGLGRITGLDPAGCFPHNTPEKVRDPDSANFVDVIHTNAARIIFELGVGTIDACGHLD	240
Qy	241	FYPNGGKHMPGCEDLITPLLKFNFNAYKK	269
Db	241	FYPNGGKHMPGCEDLITPLLKFNFNAYKK	269

## RESULT 8

ABB79040  
ID ABB79040 standard; Protein; 469 AA.

AC ABB79040:

01-AUG-2002 (first entry)

XX DE Myocastor covyous pancreatic lipase related protein precursor SEQ:4.

xx Lipase; therapeutic; pancreatic lipase related protein precursor;  
kw enzyme.

XX  
50 Myocastor covilleiXX  
PN  
IIS2002052034-A1XX  
PD  
02-MAY-2002XX  
DE  
14-DEC-2000.

XX  
DB 38-95B-2000. 300018-235075P

XX  
PA  
(CITEG /) CITEG EP Y

PA (WEBS/) WEBSTER M.  
PA (WEBS/) WEBSTER M.  
PA (WEBS/) WEBSTER M.

PA (DFRA/) DI FRANCESCO V.  
PA (DFRA/) DI FRANCESCO V.  
PA (DFRA/) DI FRANCESCO V.

XX  
BT  
Discussion of Mathematics V

XX  
1122 112207/12

[illegible]

pancreatic lipase subfamily, and is useful to provide new drug targets in the discovery of therapeutics for lipase-mediated disease -



Rat pancreatic lipase related protein precursor SEQ ID NO:6.

Lipase; therapeutic; pancreatic lipase related protein precursor; enzyme.

Rattus norvegicus.

US2002052034-A1.

02-MAY-2002.

14-DEC-2000; 2000US-0735933.

28-SEP-2000; 2000US-235925P.

(GUEG/) GUEGLER K.

(WEBB/) WEBSTER M.

(KETC/) KETCHUM K A.

(DFRA/) DI FRANCESCO V.

(BEAS/) BEASLEY E M.

Guegler K, Webster M, Ketchum KA, Di Francesco V, Beasley EM;

WPI; 2002-443697/47.

A lipase protein identified from the human genome is related to the pancreatic lipase subfamily, and is useful to provide new drug targets in the discovery of therapeutics for lipase-mediated disease

Disclosure; Fig 2; 85pp; English.

The present invention describes a human lipase protein. Molecules from the present invention can be used as models for the development of human therapeutic targets, aid in the identification of therapeutic proteins and serve as targets for the development of human therapeutic agents that modulate lipase activity in cells and tissues that express lipase. The present sequence represents a pancreatic lipase related protein precursor isolated from rat, which is given in comparison with the human lipase in the exemplification of the present invention.

Sequence 467 AA;

Query Match 46.4%; Score 1172; DB 23; Length 467;  
 Best Local Similarity 49.3%; Pred. No. 2.8e-109;  
 Matches 232; Conservative 73; Mismatches 158; Indels 8; Gaps 6;  
 1 MLCIWIIVAFLLFGTSGKEVCYVERLGCFCGDPWTRTFSTELVGLPWSPEKINTRELLYT 60  
 1 MTLTWVSLFILGAQGGKEVCYVDNLCGTSDAEPWAGTAIRPLKLLPWSPEKINTRELLYT 60  
 61 IHNPNAYQEISAVNSSTIOASYFGTDKTRINIAGW--KTDGKQWQDMCNVLLQLEDINC 118  
 61 NENPTAFQTLQSLDPLTIGASNFQVARKTRTFIHGFIDKGSWVTDCKMLFVEEVNC 120  
 119 INLDWINGSR-EYIHAVNNLRVVGAVYFIDVLMKKFEYSKVLHIGSLGHAHAGEA 177  
 121 ICVDWKKGSQATYTQANNRVVVGAVQVQMDILVKNYSYSPSKVHLIGHSLGAHAGEA 180  
 178 GSRIPLGLGRITGLDPAGPFPHNTPKPEVRLDPSDANFVDVIHTNAARILFELGVGTIDACG 237  
 181 GSRTPLGLGRITGLDPVEANFECTPEEVRLDPSDADFVDVIHTDAAPLIFLFGTQMGS 240  
 238 HLDVFNKGKMPGC-EDLIPLLFKNFNAYKEMASFFDCNHRASQYFAESILNPDAF 296  
 241 HLDVFPNGGQSMPCGCKKALQIV--DIDGWSGTRDFVACNHLRSYKYLESLNPDSG 298  
 297 IAYPCRSYTSFKAGNCFPCSGEGCTMGHFAHDFHFKNMKNTGSHYFLNTGSLSPFARWR 356  
 299 AAYPCASYKDPESNCKFCPCDQCGPCQMGHYADKAGKS-GDEPQKFFLNTGEAKNFARWR 357  
 357 HKLSVKLSGSEVTVQTVFLRVGAGTGTGEPFAYVSGKLEPQMTYTKLIDAVNVGNITSV 416  
 358 YVSVLSILGRWVT-QGVKVALFGSGKNTROYDIPRGIIKPGATHSSSEFDKLDVGTIEKV 416

QY 417 QFIWKHLFEDSQKCGAEMVINTSGKYGYKSTFCSDIMGPNIQLNKKPC 467  
 Db 417 KFLWNNQVINSPFKVGAAKITVQGBERTYFNFCSEETVREDTLLTLPLC 467

RESULT 11

ABB06317

ID ABB06317 standard; Protein; 467 AA.

XX ABB06317;

DT 27-MAY-2002 (first entry)

XX Plrpl amino acid sequence.

XX Human; phospholipase A1; PLA1; enzyme; lipase; phosphatidic acid;  
 KW cytosolic; cellular regulation; transmembrane protein; cancer;  
 KW lipase regulation disorder.

XX Unidentified.

XX WO200202762-A1.

XX 10-JAN-2002.

XX 03-JUL-2000; 2000WO-JP04441.

XX 03-JUL-2000; 2000WO-JP04441.

XX (MOCH) MOCHIDA PHARM CO LTD.

XX Inoue K, Arai H, Aoki J;

XX WPI; 2002-188424/24.

XX New human phospholipase PLA1 which specifically cleaves phosphatidic  
 PT acid, useful in the diagnosis, treatment and prevention of diseases  
 associated with lipase regulation disorders, including cancer

XX Example; Fig 3; 68pp; Japanese.

XX The present invention describes human phospholipase A1 (PLA1) which  
 specifically cleaves phosphatidic acid (PA). PLA1 has cytostatic  
 activity. Phosphatidic acid is a messenger for cellular regulation  
 and the lipase is a transmembrane protein involved in its control.  
 PLA1 can be used in the diagnosis, treatment and prevention of  
 diseases associated with lipase regulation disorders, including cancer.  
 The present sequence represents an amino acid sequence given in  
 comparison with the human PLA1 amino acid sequence, in an example from  
 the present invention.

XX Sequence 467 AA;

Query Match 46.0%; Score 1162; DB 23; Length 467;  
 Best Local Similarity 48.6%; Pred. No. 2.8e-107;  
 Matches 229; Conservative 75; Mismatches 159; Indels 8; Gaps 6;

QY 1 MLCIWIIVAFLLFGTSGKEVCYVERLGCFCGDPWTRTFSTELVGLPWSPEKINTRELLYT 60  
 Db 1 MLIWTITLFLGAQGGKEVCYEDLGCFSDEPWGTAIRPLKLLPWSPEKINTRELLYT 60

QY 61 IHNPNAYQEISAVNSSTIOASYFGTDKTRINIAGW--KTDGKQWQDMCNVLLQLEDINC 118  
 Db 61 NENPNFQILLSDPSSTIASNFQMDKTRTFIHGFIDKGSWVTDCKMLFVEEVNC 120

QY 119 INLDWINGSR-EYIHAVNNLRVVGAVYFIDVLMKKFEYSKVLHIGSLGHAHAGEA 177  
 Db 121 ICVDWKKGSQATYTQANNRVVVGAVQVQMDILVKNYSYSPSKVHLIGHSLGAHAGEA 180

QY 178 GSRIPLGLGRITGLDPAGPFPHNTPKPEVRLDPSDANFVDVIHTNAARILFELGVGTIDACG 237  
 Db 181 GSKTPGLSRITGLDPVEASPESTPEEVRLDPSDADFVDVIHTDAAPLIFLFGTQMGS 240



(NOVO) NOVO-NORDISK AS.

Gormsen E, Hjorth SA, Thim L, Verger R, Woldike HF;

WPI; 1993-036374/04.  
N-PSDB; AAQ35058.

Guinea pig and human pancreatic lipase(s) and variant(s) - used  
in detergents or as digestive enzymes for patients with cystic  
fibrosis and chronic pancreatitis

Claim 1; Page 38-40; 65pp; English.

A cDNA library was constructed from mRNA isolated from homogenised  
guinea pig pancreas. The library was transformed into competent SC51  
cells. Recombinant clones were transferred to nitrocellulose filters  
and probed with degenerate probe pools based on the amino acid  
sequences of peptide fragments from purified GPL (see AAQ35060 and  
AAQ35061). Four clones were positive to both probe pools and were  
selected as putative GPL clones. All four were confirmed as GPL cDNA  
clones by sequencing. The GPL sequence shows a high degree of  
homology to other mammalian pancreatic lipases except for a distinct  
discrepancy in the central part of the gene. This region encompasses  
the 23 amino acid "loop structure" of human pancreatic lipase which  
is substituted by a 5 amino acid "mini-loop" in GPL (see AAQ35059 for  
coding HPL sequence). GPL encoded by this cDNA sequence can be used  
in detergent compositions or to treat lipase deficiency. Also  
claimed are C-terminally truncated forms of GPL and variants  
carrying an insertional substitution, pref. of the HPL loop  
structure in place of the GPL mini-loop and/or substitution of  
Cys103 or Cys105, pref. by Ser or Thr and further by substitution  
of Ser154 by Thr and/or Asp178 by Glu.  
(Updated on 25-MAR-2003 to correct PN field.)

Sequence 452 AA;

Query Match 45.7%; Score 1156; DB 14; Length 452;  
Local Similarity 48.0%; Pred. No. 1.1e-106;  
Matches 22; Conservative 64; Mismatches 154; Indels 28; Gaps 6;  
1 MLCIIVAFVFFGTSRGEVCYERLGCPCDGLPWTFTSTELVGLPWPSPKINTRFLLYT 60  
2 MLFAWTGLLLATVRGAECYVSHLGCPSDEXPWAGTSORPIKSLPSPDKINTRFLLYT 61  
61 IHNPNAYQISAVNSSTIOASYGTDKTRINAGWKTGK--WORDMNCVLLQLEDINC 118  
62 NENQNSYQLTATDIATIKASNFNLNRKTRFIHGTDSGENSLSDMKMFMQVEKVCNC 121  
119 INLDWINGSR-EYIHAVNNLRVVGAEVAYFIDVLMKKPEYSPSKVHLGHSLGAHLAGEA 177  
122 ICVDWKGSRAQYSQASQNRVVGAEVAVLVQVLSLNYAPENVHIGHSIGAHTAGEA 181  
178 GSRIPGL-GRITGLDPAGPFHTPKEVRLDPSDANFVDTNARILFELGVGTIDAC 236  
182 GKRLNGLVGRITGLDPAEPYFQDTPKEVRLDPSDAKFDVDTSDTSPILSLGFGMSQKV 241  
237 GHLDYFNGKGMPPGCEDLITPLLKFNFNAYKEMASFFDCNHRASYOFYAESILNPDAP 296  
242 GHWDFFNGKGMPPGCKTGIS-----CNHRSISYYHSSILNPGPF 282  
297 IAYPCRSYTFKAGNCFPGKCEGPTMGHFAFRPHFNKMTNG--SHYFLNTGSLSPPAR 354  
283 LGYPCASYDFEBSGCPFCPAKGCPRMGHFAQY---PGKTNAVEQTFFLTATGASDNFTR 339  
355 WRHKLVSKLSSGVSTQCTVFLRVGGAGTGEPAIVSGKLEFGMYTKLIDADVNGVNI 414  
340 WRKYVTYVLSGKDPGSGNINVALLGKNGNSAQVQFKGLTKFDASYTNSIDVELNVTIQ 399  
415 SVQFIWKHILFEDSONKLGAEVINTSGKYGYKSTFCSDQINGPNILQKRPC 467  
400 KVTFLWKRSGISVSKPMGASRITVQSGKDGTKYNFCSSDIQVENVQTLSPC 452

RESULT 14

ABG16896  
ID ABG16896 standard; Protein; 469 AA.  
XX  
AC ABG16896;  
XX  
DT 18-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #16887.  
XX  
KW Human; chromosome mapping; Gene mapping; Gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX  
PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
PA (HYSB-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI: 2001-639362/73.  
DR N-PSDB; AAS81083.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
XX Claim 20; SEQ ID No 47255; 103pp; English.  
PS  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of DNA and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 469 AA;  
Query Match 45.6%; Score 1152; DB 22; Length 469;  
Best Local Similarity 47.8%; Pred. No. 2.9e-106;  
Matches 225; Conservative 68; Mismatches 172; Indels 6; Gaps 5;  
QY 1 MLCIIVAFVFFGTSRGEVCYERLGCPCDGLPWTFTSTELVGLPWPSPKINTRFLLYT 60  
DB 1 MLFPWTGLLLATVRGEVCYVSHLGCPSDEXPWAGTSORPIKSLPSPDKINTRFLLYT 60  
QY 61 IHNPNAYQISAVNSSTIOASYGTDKTRINAGW--KTDGKWORDMNCVLLQLEDINC 118  
DB 61 NENPNFQITGTETPTIEASNFQLDKTRFIHIGFLDKAEDSWPSDMCKMFMQVEKVCNC 120



Query Match	45.6%;	Score 1152;	DB 23;	Length 469;
Best Local Similarity	47.8%;	Pred. No. 2.9e-106;		
Matches	225;	Conservative 68;	Mismatches 172;	Indels 6; Gaps 5;
Qy	1	MLGIWIVAFLLFFCTSGKVCYERLCGFCGDXGLPWITFTSTELVGLPWSPEKINTRLLYLT	60	
Db	1	MLSPWTLGLLLLVATVSGKVCYQGLGCFSDERPWAGTLQRPVKLLPWSSEDDITRELLYT	60	
Qy	61	IHPNAYQEISAVNSSTIQASYTGTDKITRIINAGW--KTDDGKWQZDMCNVLLQLEDINC	118	
Db	61	NENPNNFQILITGTEPDTIEASNFOLDKRTRFIHGFLDKAEDSWSPDMCKMPEVEKVCNC	120	
Qy	119	INLDWTNGSRE-YIHAVNNLRVVGAVAYFDVLMKKPFYSPSKVHLIHGSLGAHLAGA	177	
Db	121	ICVDWRHGSAMTQAVQINRVVGAETAFLLIQALSTQLGYSLUEDVHIVIGSLGAHTAAEA	180	
Qy	178	GSRIPI- LGRITGLDPAGPFFHNTPKVELRDPDSANFVDVIHTNAARIILFELGVGTIDAC	236	
Db	181	GRLLGGRVGRITGLDPAGPCFQDEPEVEVRDPDSAVFVDVIHTDSSPIVPSLPGMGSKVQ	240	
Qy	237	GHLDYFVNGKGKMPGCCEDIITPLKFNFNAYKEMASFPDCKHARSYQVPAESILNPDAF	296	
Db	241	GHLDYFVNGKGKMPGCKKNVLSTIT-DIDGIEWEGIGGVSCNHLRSFEYISSVLNPDDG	299	
Qy	297	IAYPCRSYTSFXAGNCFCSKGCPTMGHFADRPHFKNNKTNKSGHYFLNTGLSLSPARWR	356	
Db	300	LGYPCASYDBEQSKCFPCPAGCPKNGHYAQDFGKTSVAECT--FFLNTGESGNFTSNR	358	
Qy	357	HKLVSKLSGSEVTQGVTVFLRVGAIGKTGEFAIVSGKLEPGWYTYKLIDADVNVGNITSV	416	
Db	359	YKVSVTLSGKEKVGNIIRALGYNSNNSKQYEIFKGLSKLPDASHATCAIDVDENVGKIQV	418	
Qy	417	QFTWKXHLFSDSONKLGAEMWINTSGKYKSTFCFSQDIMGPNILQNLRKPC	467	
Db	419	KFWANRKGINLSEPKLGASOITVOSGEDCTEYNFCSSDTEVBENVLQSLYPC	469	

Search completed: January 6, 2004, 11:21:31  
Job time : 54 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

protein - protein search, using sw model

l: January 6, 2004, 11:12:33 ; Search time 28 Seconds  
(without alignments)  
1603.956 Million cell updates/sec

t score: 2527  
ce: 1 MLGWIIVAFLEFCTSRGKEV.....STFCSQDIMGPNILQNLKPC 467

g table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

ed: 283308 seqs, 96168682 residues

number of hits satisfying chosen parameters: 283308

m DB seq length: 0

m DB seq length: 2000000000

rocessing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

se : PIR\_761\*

1: PIR1.\*

2: PIR2.\*

3: PIR3.\*

4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Score	Query Match	Length	DB ID	Description
1191	47.1	470	2 A54232	lipase, CoPL-RP2 -
1174	46.5	467	1 LIDG	triacylglycerol li
1172	46.4	473	2 S20612	triacylglycerol li
1162	46.0	467	2 A43357	pancreatic lipase-
1152	45.6	469	2 B43357	pancreatic lipase-
1147.5	45.4	457	2 I48206	triacylglycerol li
1142.5	45.2	482	1 A46696	triacylglycerol li
1129.5	44.7	482	1 A34571	triacylglycerol li
1121	44.4	434	1 A49488	triacylglycerol li
1115	44.1	465	2 C43357	triacylglycerol li
1109	43.9	461	2 S21223	triacylglycerol li
1108	43.8	465	1 S41084	triacylglycerol li
1107	43.8	465	1 JC1318	triacylglycerol li
1039.5	41.1	449	1 LIPG	triacylglycerol li
534	21.1	499	2 A28997	triacylglycerol li
503	19.9	510	2 S15993	triacylglycerol li
483	19.1	494	2 A27442	triacylglycerol li
471	18.6	465	1 A27330	lipoprotein lipase
459	18.2	478	2 S18158	lipoprotein lipase
458.5	18.1	490	1 S04331	lipoprotein lipase
453	17.9	475	1 LTHUL	lipoprotein lipase
452	17.9	478	2 S29846	lipoprotein lipase
450	17.8	475	2 JC4242	lipoprotein lipase
445.5	17.6	474	1 JH0790	lipoprotein lipase
444.5	17.6	450	1 A27053	lipoprotein lipase
442.5	17.5	474	1 A40570	lipoprotein lipase
318.5	12.6	172	2 A45658	triacylglycerol li
287	11.4	303	2 A44563	phospholipase A1
284.5	11.3	300	2 A44564	phospholipase A1

30 277 11.0 317 2 S32406 phospholipase A1 -  
31 216 8.5 439 1 VJFF1 vitellogenin 1 pre  
32 209.5 8.3 356 2 S49011 yolk protein 2 - b  
33 202 8.0 422 2 S49012 yolk protein 3 - b  
34 192 7.6 420 2 A25876 vitellogenin III p  
35 184.5 7.3 442 1 VJRF2 vitellogenin II pr  
36 183 7.2 422 2 S22888 vitellogenin 2 - M  
37 158.5 6.3 437 2 S22889 vitellogenin 1 - M  
38 119 4.7 254 2 AC0095 probable thioester  
39 102.5 4.1 303 2 S33490 hypothetical 303aa  
40 102 4.0 1603 2 T24098 hypothetical prote  
41 97 3.8 508 1 ALMSP alpha-amyliase (EC  
42 97 3.8 508 1 ALMSPA alpha-amyliase (EC  
43 96.5 3.8 989 2 I56333 apolipoprotein B -  
44 96 3.8 1140 2 JC7152 UV-damaged DNA-bin  
45 95.5 3.8 499 2 C75251 phosphopyruvate hy

ALIGNMENTS

RESULT 1

A54232  
lipase, CoPL-RP2 - nutria  
C:Species: Myocastor coypus (nutria, coypu)  
C:Date: 22-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 16-Feb-1997  
C:Accession: A54232  
R:Thirstrup, K.; Verger, R.; Carriere, F.  
Biochemistry 33, 2748-2756, 1994  
A:Title: Evidence for a pancreatic lipase subfamily with new kinetic properties.  
A:Reference number: A54232; MUID:94176463; PMID:9130186  
A:Accession: A54232  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-470 <THI>  
A:Experimental source: pancreas  
A:Note: sequence extracted from NCBI backbone (NCBIP:144530)  
C:Superfamily: triacylglycerol lipase

Query Match 47.1%; Score 1191; DB 2; Length 470;  
Best Local Similarity 49.8%; Pred. No. 2.5e-88;  
Matches 236; Conservative 63; Mismatches 163; Indels 12; Gaps 7;  
Qy 1 MLGWIIVAFLEFCTSRGKEVYERLGCFCFKDGLPWRTRTFSTBLVGLPWPSPKINTRFLLYT 60  
Db 2 MLFVWTTGLLLATARGNEVCVSHLGCFSDEKPMAGTLQRPVKSILPASPEINTRFLLYT 61  
Qy 61 IHNPNAYQISAVNSSTIOASYFGTDKJTRINIAWKTDGK--WORDMNVLLQLEDINC 118  
Db 62 NENPNYQITATDPTATIKASNFNLHRTFRFVHGFIDNGKDWLTD:CKRMFQVEKVC 121  
Qy 119 INLDWINGRE-YIHAVNLRVVGAEVAVFDVLMKPEYSPKVHLHGSLGAHAGEA 177  
Db 122 ICVDWOGGSLAYSAQVQNIHVGAELVQLVLSDLQGYKPGNVHMGHSLGAHTAAEA 181  
Qy 178 GSRIPGL-GRITGLDPAGFFHNTPKVRLPDSANFVDVHTHTNAARILFELGVGTIDAC 236  
Db 182 GRRLKLVGRITGLDPAEPFCQDTPPEVRLPDSANFVDVHTHTDIAPISFGFGMSQKV 241  
Qy 237 GHLDYFPGNGKHPGCE-DLITPLKFNPAKXKEMASFFPDONHARSYQFYAESLNPDA 295  
Db 242 GHMDFFPGNGKHPGCEKNIIITIV--DVNGFLEGITSLAACNHNRSYQYSSSLNPDG 299  
Qy 296 FIAYPCRSYSTSKAGNCFPCSKEGCPTMGHFAFRPHFKNMKTNG--SHYFLNTGSLSPFA 353  
Db 300 FLGYPCASYEEFKDGCFCPCPAEGCPCPKGHVADQF--QGKANGVEKTYFLNTGSDSNFP 356  
Qy 354 RWRHKLKVLKSESEVTOGTVFRVCGALGKTGEFAIVSGKLEBPGMTYTKLIDADVNCNI 413  
Db 357 RWRKYKSVTLSEKELSGDIKIALFGRNGSKQVEIFKGSILKPDARYTHD:DVLDNNGEI 416  
Qy 414 TSQPTWKKHLFEDSONKLGASWINTSGKYGKSTFCSDIMGNINQNLKPC 467

417 QKVKFLWHNNINLQPKLGASQITVQSGEYGTKNFCSSNTVQSDVLQSLSPC 470

T 2

Ylglycerol lipase (EC 3.1.1.3) precursor, pancreatic - dog  
Enzyme names: triacylglycerol acylhydrolase  
Species: Canis lupus familiaris (dog)  
Date: 30-Sep-1987 #sequence revision 30-Sep-1987 #text\_change 31-Mar-2000  
Accession: B24392; A32615; E30034  
Molecule type: mRNA  
Residues: 1-430-437, 1986  
Cross-references: EMBL:X61925; NID:956599; PIDN:CAA43927.1; PID:956600  
Superfamily: triacylglycerol lipase  
Keywords: carboxylic ester hydrolase

Query Match 46.4%; Score 1172; DB 2; Length 473;  
Best Local Similarity 49.3%; Pred. No. 8.5e-87;  
Matches 232; Conservative 73; Mismatches 158; Indels 8; Gaps 6;

QY 1 MLGIWIVAFPGTSRGKVCYERLGCFOGLPWTFTSTELVGLPWSPEKINTRFLLYT 60

DB 1 MLGIWIVAFPGTSRGKVCYERLGCFOGLPWTFTSTELVGLPWSPEKINTRFLLYT 60

QY 61 IHNPNAYQISAVNSTIOASYFGTDKTRINIAW--KTDGKQWDMCNVLLQLEIDNC 118

DB 61 IHNPNAYQISAVNSTIOASYFGTDKTRINIAW--KTDGKQWDMCNVLLQLEIDNC 118

QY 119 INLDWINGSR-EYIHAVNNLRVVGAEVAYFDVLMKKFEYSPSKVHLIGHSLGAHLAGEA 177

DB 119 INLDWINGSR-EYIHAVNNLRVVGAEVAYFDVLMKKFEYSPSKVHLIGHSLGAHLAGEA 177

QY 121 ICVDWKKGSGTQYTOAANNRVVGAQVACMLSMLSANYSYSPSQVQLIGHSLGAHVAGEA 180

DB 121 ICVDWKKGSGTQYTOAANNRVVGAQVACMLSMLSANYSYSPSQVQLIGHSLGAHVAGEA 180

QY 178 GSRTPGLGRITGLDPAGPFFHNTPEKVRLDPSDANFVDVHTNAARILFELGVGTIDACG 237

DB 178 GSRTPGLGRITGLDPAGPFFHNTPEKVRLDPSDANFVDVHTNAARILFELGVGTIDACG 237

QY 181 GSRTPGLGRITGLDPAGPFFHNTPEKVRLDPSDANFVDVHTNAARILFELGVGTIDACG 240

DB 181 GSRTPGLGRITGLDPAGPFFHNTPEKVRLDPSDANFVDVHTNAARILFELGVGTIDACG 240

QY 238 HLDYPNGGKHPGC-EDLITPLKFNFNAYKEMASFFDCNARSYQFYAESILNPDAP 296

DB 238 HLDYPNGGKHPGC-EDLITPLKFNFNAYKEMASFFDCNARSYQFYAESILNPDAP 296

QY 241 HLDYPNGGKHPGC-EDLITPLKFNFNAYKEMASFFDCNARSYQFYAESILNPDAP 298

DB 241 HLDYPNGGKHPGC-EDLITPLKFNFNAYKEMASFFDCNARSYQFYAESILNPDAP 298

QY 297 IAYPCRSYTSFKAGNCFKCGEPTMGHFAHDFHFKMKTNGSHYFLNTGSLSPFARWR 356

DB 297 IAYPCRSYTSFKAGNCFKCGEPTMGHFAHDFHFKMKTNGSHYFLNTGSLSPFARWR 356

QY 299 ASVPCASYAFESNCKFCPCDQCGPQMGHYADKFAVKT-SDETOKYFLNTGDSNPFARWR 357

DB 299 ASVPCASYAFESNCKFCPCDQCGPQMGHYADKFAVKT-SDETOKYFLNTGDSNPFARWR 357

QY 357 HKLSVKLSGSEVTGQTVFLRVGGAIGKTGEFAIVSGKLEPGMTYTKLIDADVNVGNITSV 416

DB 357 HKLSVKLSGSEVTGQTVFLRVGGAIGKTGEFAIVSGKLEPGMTYTKLIDADVNVGNITSV 416

QY 417 QFTWKHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDQIMGNILQNLKPC 467

DB 417 QFTWKHLFEDSQNKLGAEMVINTSGKYGYKSTFCSDQIMGNILQNLKPC 467

QY 417 KFLMNNVNVNPFPPKVGAAKITVQKGEKTVHSCSESTVEDTLLTTPC 467

DB 417 KFLMNNVNVNPFPPKVGAAKITVQKGEKTVHSCSESTVEDTLLTTPC 467

RESULT 4

A43357

pancreatic lipase-related protein 1 precursor - human

Species: Homo sapiens (man)

Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text\_change 31-Mar-2000

Accession: A43357

Molecule type: mRNA

Residues: 1-467 <GIL>

Cross-references: GB:M93283; NID:g187229; PIDN:AAA59532.1; PID:g187230

Note: sequence extracted from NCBI backbone (NCBIN:110821, NCBI:110822)

Genetics: GDB:PLNLRP1; PLRP1

Cross-references: GDB:134816

Superfamily: triacylglycerol lipase

F1-17/Domain: signal sequence #status predicted <SIG>

F18-467/Product: pancreatic lipase-related protein 1 #status predicted <MAT>

T 3

2

17 Match 46.0%; Score 1162; DB 2; Length 467;  
 Local Similarity 48.8%; Pred. No. 5.4e-86;  
 Cons 229; Conservative 75; Mismatches 159; Indels 8; Gaps 6;  
 1 MLGIWIVAFLEFGTSRGKEVCYERLGCFCGDLGTPWTRTFSTELVGLPWSPEKINTRELLYT 60  
 1 MLIFWITLFLGAAKKEVCYEDLGCSDTEPGWGTARLPKILPWSPEKIGTRFLYT 60  
 61 IHNPNAYQISAVNSSTIOASYFGTDKTRINIAW--KTDGKWORDMNCVLLQLEDINC 118  
 61 NENPNFQILLSDPSTIEASNFQMDRTRFIHGFIDKGDSEWVDMCKLFEVEVNC 120  
 119 INLDWINGSR-EYIHAVNNLRVVGAEVAFIDVLMKKFEYSPSKVHLIGHSLGAHAGEA 177  
 121 ICVDWKKGQAYTOANNRVVVGAAQMLDILTEYSYPPSKVHLIGHSLGAHAGEA 180  
 178 GSRIPGLRITGLDPAGPPFHNTPKEVRLDPSDANFVDVHTNAARILFELGVGTIDACG 237  
 181 GSKTPGLSRITGLDPVEASFESTPEVRLDPSDADFVDVHTDAAPLIPFLGFGTNOQM 240  
 238 HLDFFPNGKHPGCG-EDLITPLKFNFNAYKEMASFFDCNHARSYQFYAESILNPDFA 296  
 241 HLDFFPNGGSMGCKKXALSOIV--DLDGWAGTRDFVACNHLRSYKYLESLINPDGF 298  
 297 IAYPCRSYTSFKAGNCFCKSGEGCPTMGHFAFDRHFNKMTKNGSHYFLNTGSLSPPARWR 356  
 299 AAYPCTSYKSFESDKCFPCDQCGQMGHYADKFAGRTSEQ-QKFFLNTGEASNFARWR 357  
 357 HKLSVKLSGSEVTQGTFLRVGAGIKGTGEFAIVSGKLEPGMTYTKLIDADVNGNITSV 416  
 358 YGVSTLISRTAT-GQIKVALFGNKGNTHOYSIFRGLKPGSTHSYEFDAKLDVGTIEK 416  
 417 QRIWKXHLPEDSQNKLGAEWINTSGKYKSTFCSDQIMGNILQNLKPC 467  
 417 KLMNNVNLPLRVGATKITVQGEKTYNFCSEDTVREDTLLTLTPC 467  
 5  
 satic lipase-related protein 2 precursor - human  
 Species: Homo sapiens (man)  
 Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 31-Mar-2000  
 Accession: B43357  
 Ler, T.; Buchwald, P.; Blum-Kaelin, D.; Hunziker, W.  
 J. Biol. Chem. 267, 16509-16516, 1992  
 Note: Two novel human pancreatic lipase related proteins, hPLRP1 and hPLRP2. Differ-  
 ence number: A43357; MUID:92355622; PMID:1379598  
 Accession: B43357  
 Molecule type: mRNA  
 Features: 1-469 <GL>  
 References: GB:M93284; NID:g187231; PIDN:AAAS9533.1; PID:g187232  
 Date: sequence extracted from NCBI backbone (NCBIN:110824, NCBI:P.110825)  
 Notes:  
 1 GDB:PLNLRP2; PLNRP2  
 References: GDB:134817  
 Family: triacylglycerol lipase  
 Domain: signal sequence #status predicted <SIG>  
 167/Product: pancreatic lipase-related protein 2 #status predicted <MAT>  
 Query Match 45.6%; Score 1152; DB 2; Length 469;  
 Best Local Similarity 47.8%; Pred. No. 3.5e-85;  
 Cons 225; Conservative 68; Mismatches 172; Indels 6; Gaps 5;  
 1 MLGIWIVAFLEFGTSRGKEVCYERLGCFCGDLGTPWTRTFSTELVGLPWSPEKINTRELLYT 60  
 1 MLPPWTLGLLLATVRGKEVCYGLGCFSDKPKWAGTLQRPVKLLPWSPEKIDITRELLYT 60  
 61 IHNPNAYQISAVNSSTIOASYFGTDKTRINIAW--KTDGKWORDMNCVLLQLEDINC 118  
 61 NENPNFQILLTGTEDTIEASNFQMDRTRFIHGFIDKAEWSDMPDMCKMFEVEKVC 120  
 119 INLDWINGSR-EYIHAVNNLRVVGAEVAFIDVLMKKFEYSPSKVHLIGHSLGAHAGEA 177

121 ICVDWRHGRAMTYTQAVQNIWVGAEAFILQALSTQLGYSLSDVHVHIGSLGAHTAAEA 180  
 178 GSRIPG-LRITGLDPAGPPFHNTPKEVRLDPSDANFVDVHTNAARILFELGVGTIDAC 236  
 181 GRLGRVGRITGLDPAGCFQDPEVRLDPSDAVFDVHTDSSPIVPSLGFMSQKV 240  
 237 GHLDFYPNGKHPGCEDELIITPLKFNFNAYKEMASFFDCNHARSYQFYAESILNPDFA 296  
 241 GHLDFPFGKHPGCKKXALSTIT-DIDGIWEGIGGFVSCNHLRSFEYSSVNLNPDGF 299  
 297 IAYPCRSYTSFKAGNCFCKSGEGCPTMGHFAFDRHFNKMTKNGSHYFLNTGSLSPPARWR 356  
 300 LGYPCASVDEFOESKCFPCPAEGCPKMGHYADQPKGTSAVEQT-FFELNTGESGNFTSWR 358  
 357 HKLSVKLSGSEVTQGTFLRVGAGIKGTGEFAIVSGKLEPGMTYTKLIDADVNGNITSV 416  
 359 YKVSVTLSGKGVNGYRIALYGSNENSKQYIEFKGSLKPDASHTCALDIDVFNVGKIQV 418  
 417 QRIWKXHLPEDSQNKLGAEWINTSGKYKSTFCSDQIMGNILQNLKPC 467  
 419 KFLMNRGINLSEPKLGASQITVQSGEDGTEYNFCSSDTVEENVLSLYPC 469  
 RESULT 6  
 I48206  
 triacylglycerol lipase (EC 3.1.1.3) - nutria (fragment)  
 Species: Myocastor coypus (nutria, coypu)  
 Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 18-Jun-1999  
 Accession: I48206  
 R.Thiistrup, K.; Carriere, F.; Hjorth, S.A.; Raamussen, P.B.; Ladefoged, C.  
 Eur. J. Biochem. 227, 186-193, 1995  
 Note: Cloning and expression in insect cells of two pancreatic lipases and a procolip-  
 A:Reference number: I48204; MUID:95154288; PMID:7851384  
 A:Accession: I48206  
 A:Status: preliminary; translated from GB/EMBL/DBD  
 A:Molecule type: mRNA  
 A:Residues: 1-457 <RES>  
 A:Cross-references: EMBL:X82999; NID:g599870; PIDN:CAA58120.1; PID:g599871  
 C:Superfamily: triacylglycerol lipase  
 C:Keywords: carboxylic ester hydrolase  
 Query Match 45.4%; Score 1147.5; DB 2; Length 457;  
 Best Local Similarity 49.9%; Pred. No. 7.8e-85;  
 Matches 232; Conservative 69; Mismatches 149; Indels 15; Gaps 10;  
 10 LFFGTSRGKEVCYERLGCFCGDLGTPWTRTFSTELVGLPWSPEKINTRELLYTINPNAYOE 69  
 1 LLLGAVAGSEVCYDRLCGCFSDSPWAGIVERPLKLVFWSPTINTRELLYTNESPNNYQI 60  
 70 ISAVNSSTIOASYFGTDKTRINIAW--KTDGKWORDMNCVLLQLEDINCINLDWINGS 127  
 61 VTA-DSSTIRSSNPTDRKTRFIHGYIDKGENWLANMCEALLQVESVNCICVDWKG 119  
 128 RE-YIHAVNNLRVVGAEVAFIDVLMKKFEYSPSKVHLIGHSLGAHAGEASRIPG-LG 185  
 120 RALYQATQNIWVGAEVAFVFDALQSLGSPSNVHIIHSLGSHVAGEAGRTNGNIG 179  
 186 RITGLDPAGPPFHNTPKEVRLDPSDANFVDVHTNAARILFELGVGTIDACGHLDFPNG 245  
 180 RITGLDPAPCFQGTTELPEVRLDPSDAQFVDVHTDGAPIIPNLGFMGQTVGHLDFFPNG 239  
 246 GKHPGCE-DELIITPLKFNFNAYKEMASFFDCNHARSYQFYAESILNPDFAIYPCRSY 304  
 240 GVEMPGCKNIISQIV--DINGIWEGRDFAACNHLRSYKYVDSILNPTGAGFSCSY 297  
 305 TSFKAGNCFCKSGEGCPTMGHFAFDRHFNKMTKNG--SHYFLNTGSLSPPARWRKLSVK 362  
 298 NTSSNNCFPCASGCGPQMGHYADRP--SGKTNELFOQFYFLNTGDASNFSRWYQIAVT 354  
 363 LSGSEVTQGTFLRVGAGIKGTGEFAIVSGKLEPGMTYTKLIDADVNGNITSVQFIWCK 422  
 355 LSGRKVT-GHVLVSLVSGSGTQKQYIYKSLQPGTSYNQIDSDVDVDGIEKVFYIN 413

```

423 HLFEDSQNKLGAEMVINTSGYKSTFCSDQIMGNILQNLKPC 467
      : : : : : : : : : : : : : : : : : : : : : :
414 NIINPTLPKVGASS-IQWRNDRGVNFQSDTVREDILLTLTPC 457
      : : : : : : : : : : : : : : : : : : : : : :

>T 7
>C: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 31-Mar-2000
>Accession: A46696
>Molecule type: mRNA
>Residues: 1-482 <WTS>
>Cross-references: GB:W30687
>Note: the authors translated the codon CGA for residue 362 as Gly
>Superfamily: triacylglycerol lipase
>Keywords: carboxylic ester hydrolase
>F1-30/Domain: signal sequence #status predicted <SIG>
>F184,208,295/Active site: Ser, Asp, His #status predicted

Query Match 44.7%; Score 1129.5; DB 1; Length 482;
Best Local Similarity 48.0%; Pred. No. 2.4e-83;
Matches 229; Conservative 63; Mismatches 172; Indels 7; Gaps 6;

Qy 1 MLGIWIVAFPFSTGRKEVCYERLGCDFKGLPWTRTFSTELVGLPWSPEKINTRELLYT 60
Db 15 MLLCWIVS-LLLATVGGKEVCYHLCFSDKPKWAGMIQRFSKIFPWSPEIDTRFLLYT 73
Qy 61 IHNPNAYQISAVNSSTIQASYFGTDKITRINIAGWKTDGK--WORDMNCVLLQLEDINC 118
Db 74 NENPNYQISATDPTATNASNFQDRKTRFIIHGFIDKBEGWLLDMCKMFOVEKYNV 133
Qy 119 INLDWNGSR-EYIHAVNNLRVVGAEVAFIDVLMKKPEYSPSKVHLIGHSLGAHLAGEA 177
Db 134 ICVDWRGSRTEVTOASYNTRVVGAEIAFLVQLVSTENGYSPENVHLLIGHSLGHVGEA 193
Qy 178 GSRIPG-LGRITGLDPAGPFPHTPKEVRLDPSDANFVDVIHTNAARILFELGVGTIDAC 236
Db 194 GRLEGHVGRITGLDPAEPFCQGLPEEVRDPSDAMFVDVIHTDSAPITPYLFGMSQKV 253
Qy 237 GHLDYPNGGKHPGCCEDLITPLKFNFNAYKEMASFFDCNHARSYCFYAESILNPDAP 296
Db 254 GHLDFFNGGKHPGCCQKNILSTI-VDINGIEGTQNFVACNHLRSYKYASSILNPDGP 312
Qy 297 IAYPCRYTSFKAGNCFKSCGPTMGHFAFRPHKNNKMGSHYFINTGSLSPFFARWR 356
Db 313 LGYPCSSYEFQNDCCFPCEGCPKMGHYADQFEGKTATVEQTVV-LNTGDSGNFTWR 371
Qy 357 HKLSVKLSGSEVTQGVFLRVGAIGKTEGPAIVSGKLEPGMTYTKLIDADVNGNITSV 416
Db 372 YKVSVTLSGAKLSGILVALYGNNGSKQYEFKGLQPEARVIRIDVDVNGEIQKV 431
Qy 417 QFTWKHLFEDSQNKLGAEMVINTSGYKSTFCSDQIMGNILQNLKPC 467
Db 432 KFLWNNKVINLFRPTLGASQITVQSGVDGKEYNFCSSDTVREDVLQSLVPC 482

RESULT 9
A49488
triacylglycerol lipase (EC 3.1.1.3), pancreatic - guinea pig
C:Species: Cavia porcellus (guinea pig)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 31-Mar-2000
C:Accession: A49488
R:Hyorth, A.; Carriere, F.; Cudrey, C.; Woldike, H.; Boel, E.; Lawson, D.M.; Ferrato, F.,
Biochemistry 32, 4702-4707, 1993
A:Title: A structural domain (the lid) found in pancreatic lipases is absent in the guinea
A:Reference number: A49488; MUID:93257433; PMID:8490016
A:Accession: A49488
A:Status: preliminary; translation not shown; not compared with conceptual translation
A:Molecule type: mRNA; protein
A:Residues: 1-434 <HJO>
A:Experimental source: pancreas
A:Note: sequence extracted from NCBI backbone (NCBIP:132521)
C:Superfamily: triacylglycerol lipase
C:Keywords: carboxylic ester hydrolase
F:154,178,247/Active site: Ser, Asp, His #status predicted

```

>T 8

>C: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 31-Mar-2000

>Accession: A34671

Page 4

ry Match 44.4%; Score 1121; DB 1; Length 434;  
 t Local Similarity 48.4%; Pred. No. 1e-82;  
 ches 220; Conservative 63; Mismatches 144; Indels 28; Gaps 6;

19 EVCYERLGCDFKGLPWTFTSTFELVGLPWSPEKINTRELLYTINPNAYQEISAVNSSTI 78  
 2 EVCYSHLGCDFDEKFWAGTSQRPISKLSDFPKKINTRELLYTINENQNSYQLITADIATI 61  
 79 QASYFTGDKITRINIAWKTDGK--WORDMGNVLLQLEDINCINLNDWINGSR-EYIHAVN 135  
 62 KASNFNLNKRFTFIHGTDSGNSWLSMDCKMFWQVKVNCICVDWKGSKQAYSOASQ 121  
 136 NRVVGAEVAYFIDVLMKKFVSPSKVHLGHSIGAHLAGAGSRIPL-GRITGLDPAG 194  
 122 NRVVGAEVAYLVQVLSLNAVAPENVHIIHSLGANTAGAGKRLNGLVGRITGLDPAE 181  
 195 PFPHNTPEVRLLDPSDANFVDVHTNAARILFELGVGTIDACGHLDFVPGKGHEMCCED 254  
 182 PFQDTPEVRLLDPSDAKFDVHTDISPLPSJGFGMSQKVGHMDFPFGKMDPCCKT 241  
 255 LITPLKLFNFNAYKEMASFFDCNHARSYQFAYESILNPDATFAYPCRSYTSFKAGNCFF 314  
 242 GIS-----CNHRSIEYHSSILNPEGLGYPCASYDFEQSGCFF 282  
 315 CSKGCPTMGHFAFRFPAKNTNG--SHYFNTGSLSPFAPRHRKLSVKLSGEVTOGT 372  
 283 CPKGCPRKGFADQY---PGKNAVEGTFELNGASDNFTRRYKVTYVTLSEKDPGSGN 339  
 373 VFLRVGGAIGTKTGEPAIVSGKLEPOMTYTKLIDADVNGNITSVQFTKXKHLDFSONKL 432  
 340 INVALLGKNGNSAQVQVFKGLTKDPASYSIDVELANGVTLQKTFELKSGSISVSKPKM 399  
 433 GAEWINTSGKYGKYSTPCSDODMGPNILQNLKPC 467  
 400 GASRTVQSGKDGTKYCNFSSDIVQENVEQTLSPC 434

f 10  
 ,  
 /glycerol lipase (EC 3.1.1.3) precursor, pancreatic - human  
 arinate names: fatty acid ethyl ester synthase; pancreatic lipase  
 ies: Homo sapiens (man)  
 s: 30-Sep-1993 #sequence revision 30-Sep-1993 #text change 31-Mar-2000  
 ssion: C43357; A34494; S08272; A60135; S34578; A34574  
 er. T.; Buchwald, P.; Blum-Kaelin, D.; Hunziker, W.  
 l. Chem. 267, 16509-16516, 1992  
 e: Two novel human pancreatic lipase related proteins, hPLRP1 and hPLRP2. Differ  
 erence number: A43357; MUID:92355622; PMID:1379598  
 ssion: C43357  
 us: nucleic acid sequence not shown; translation not shown  
 ule type: mRNA  
 idues: 1-465 <GIL>  
 s-References: GB:M93285; NID:G190139; PIDN:AAA60129.1; PID:G190140  
 ler, F.K.; Rosenblum, J.L.; Strauss, A.W.  
 l. Chem. 264, 20042-20048, 1989  
 le: Cloning and characterization of human pancreatic lipase cDNA.  
 erence number: A34494; MUID:90062115; PMID:2479644  
 ssion: A34494  
 ule type: mRNA  
 idues: 1-465 <LOW>  
 s-References: GB:J05125; NID:G339596; PIDN:AAA36740.1; PID:G339597  
 cler, F.K.; D'Arcy, A.; Hunziker, W.  
 a 343, 771-774, 1990  
 le: Structure of human pancreatic lipase.  
 erence number: S08272; MUID:90158821; PMID:2106079  
 ssion: S08272  
 us: not compiled with conceptual translation  
 ule type: mRNA  
 idues: 17-46; 48-465 <WIN>  
 a: part of this sequence, including the amino end of the mature protein, was confir  
 'aro, A.; Bonicel, J.; Pieron, G.; Guy, O.  
 mie 63, 799-801, 1981

A>Title: Comparative studies of human and porcine pancreatic lipases: N-terminal sequence  
 A:Reference number: A60135; MUID:82069208; PMID:7306587  
 A:Accession: A60135  
 A:Molecule type: protein  
 A:Residues: 17-30, 'X', 32-33, 'X', 35-36, 'IOX', 40-41, 'X', 43-45 <DEC>  
 R:Thirstrup, K.; Carrier, F.; Hjorth, S.; Rasmussen, P.B.; Woeldike, H.; Nielsen, P.F.;  
 FBBS Lett. 327, 79-84, 1993  
 A>Title: One-step purification and characterization of human pancreatic lipase expressed  
 A:Reference number: S34578; MUID:93327921; PMID:8335100  
 A:Accession: S34578  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 17-31 <THR>  
 R:Riley, D.J.S.; Kyger, B.M.; Spillburg, C.A.; Lange, L.G.  
 Biochemistry 29, 3848-3852, 1990  
 A>Title: Pancreatic cholesteryl esterases. Purification and characterization of human par  
 A:Reference number: A34574; MUID:90283377; PMID:2354155  
 A:Accession: A34574  
 A:Molecule type: protein  
 A:Residues: 'X', 18-36 <RI>  
 A:Experimental source: pancreas  
 C:Genetics:  
 A:Gene: GDB:PNLIP  
 A:Cross-references: GDB:127916; OMIM:246600  
 A:Map position: 10q24-10q26  
 C:Superfamily: triacylglycerol lipase  
 C:Keywords: carboxylic ester hydrolase  
 F:1-16/Domain: signal sequence #status predicted <SIG>  
 F:17-465/Product: triacylglycerol lipase #status experimental <MAT>  
 Query Match 44.1%; Score 1115; DB 2; Length 455;  
 Best Local Similarity 47.8%; Pred. No. 3.4e-82;  
 Matches 226; Conservative 73; Mismatches 160; Indels 14; Gaps 10;

QY 1 MLGIWVAFLLPFGTSRGKVCYERLGCDFKGLPWTFTSTFELVGLPWSPEKINTRELLYT 60  
 DB 1 MLPLWTLIS-LLLGAVAGKVCYERLGCDFSDSPWSGITERPLHILPWSPKDVNTRLLYT 59  
 QY 61 IHNPNAYQEISAVNSSTIQASYFTGDKITRINIAAG--KTDGKQWQDMCNVLLQLEDINC 118  
 DB 60 NENPNFQEVAA--DSSSISGNSFNTKTRFTIIFGFDKGEENWLANVCNLFKVESVNC 118  
 QY 119 INLDWINGSRE-VIHAVNLRVVGAEVAYFIDVLMKKFVSPSKVHLGHSIGAHLAGEA 177  
 DB 119 ICVDNKGSGTGTQASQNIIRIVGAEVAYFVFLQSAFGYSNVEHVIHSLGHAAGEA 178  
 QY 178 GSRIIPG-LGRITGLDPAGPPFHTPTKPEVRLDPSDANFVDVHTNAARILFELGVGTIDAC 236  
 DB 179 GRRNTNGTIGRITGLDPAEPCFQGTPELVRDPSDAKFDVHTDGAIPVNLGFGMSQV 238  
 QY 237 GHLDFFYNGCKHMPGCCDLITPLLKFNFNAYKEMASFFDCNHARSYQFAYESILNPDPAF 296  
 DB 239 GHLDFFYNGGVEMPGCKKNLSOI-VDIDIGETGTRFAACNHLRSYKYITDSIVNPDGF 297  
 QY 297 IAYPCRSYTSFKAGNCFFCSKGCPTMGHFAFRFPAKNTN--GSHYFNTGSLSPFAP 354  
 DB 298 AGPFCASVNVFTANKCFPCPSGCGPQMGHYADRY---PGKTNVQKQFVLDGTGASNFAR 354  
 QY 355 WRHKLVLKLSGSEVTOGTVFLRVGGAIGTKTGEPAIVSGKLEPOMTYTKLIDADVNGNIT 414  
 DB 355 WRYKVSVYTLSSGKVT-GHILVSLFGKNGSKQYEIFKGLTKPDSTHSNEFSDVDVGDILQ 413  
 QY 415 SVQFIWKXHLFEDSQNKLGAEMVINTSGKYGKYSTPCSDODMGPNILQNLKPC 467  
 DB 414 MWKFIWNNVINPFLRVGASKII-VETNVGQFNFCSPTVREEVLLTLTPC 465

RESULT 11  
 S21223  
 triacylglycerol lipase (EC 3.1.1.3) precursor, pancreatic - horse (fragment)  
 C:Species: Equus caballus (domestic horse)  
 C:Date: 22-Nov-1993 #sequence\_revision 26-Jul-1996 #text\_change 18-Jun-1999  
 C:Accession: S21223; S21251

rfelec, B.; Foglizzo, E.; Bonicel, J.; Bougis, P.E.; Chapus, C.  
J. Biochem. 206, 279-287, 1992  
Title: Sequence of horse pancreatic lipase as determined by protein and cDNA sequencing  
Accession number: S21223; MUID:92267022; PMID:1587279

Sequence:  
S21223  
Molecule type: mRNA  
Accessions: 1-461 <KER>  
References: EMBL:X6218; NID:g1063; PIDN:CAA45961.1; PID:g1064  
Session: S21251  
Molecule type: protein  
Accessions: 13-78; 80; 97-218; 220-228; 230-277; 322-349; 384-438; 440-448; 450-461 <KEW>  
Description family: triacylglycerol lipase  
Keywords: carboxylic ester hydrolase  
Domain(s): signal sequence (fragment) #status predicted <SIG>  
Product: triacylglycerol lipase #status experimental <WAT>

Protein Match      43.9%; Score 1109; DB 2; Length 461;  
Local Similarity    48.4%; Pred. No. 1e-81;  
Conservative         66; Mismatches 165; Indels 10; Gaps 9;

5 WIVAELEFGTSGKVCYERLCGCKDGLPWTRTS TELVGLPWSPEKINFRFLLYTHNP 64  
                :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  
1 WTLS-LLLGAVVGNEVCYERLCGCSDDSPAGIVERPLKLPLEMSPEKNRFFLLYTNEPP 59  
                :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  
65 NAYQISLVNSSTIOASVFGRDKITRIINIAG--KTDPKWQRDMCNVLLOLEDINCILND 122  
                :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  
60 DNFEIIVA-DPS TIOSSNFNTGRKRTRFIHGFIDKGESWLSTMCMFKVESVNCICVD 118  
                :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  
123 WINGSR-EYIHANNLRVVGAEVAFYFDVLMKFTEYSFVKHLIGHSLGAHLAGEAGRSRI 181  
                :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  
119 WKSGRTAYSQAQQNRVTGVGAEVALYGVLQSDFSDPSNVHHIGHSGLSHAAGAAGRRT 178  
                :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  
182 PG-LGRIRITGLPGAPFPFHPTPKVRLPDSANFYDVVIHTNAARLFELGVGTIDACGHLD 240  
                :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  
179 NGAVGRITGLDPAPECFQGTELRLPDSDAQFVDVIIHTDIAPFIPNLFGMSGQTAGHLD 238  
                :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  
241 FYPNGKKMPGCCDLITPLLKENFNAYKKEMASFDCNHARSQFYAESILNPDAFIAYP 300  
                :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  
239 FPNGGGKMPCGCQNVLISOI-VDIGTIWGQTRFDAACNHLRSKYKYTTDSLNDPDGPAFS 297  
                :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  
301 CRSYTSFAKGNCCFCSEKCCTMGHFADRFHKNMKGSHGYFLTNGTLSPPARWEHKL 360  
                :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  
298 CASYSDFTANKCFCSCSEGCGQMGIADRPFGRGT-KGVQLFYLNTGDASNFAWRMYRD 356  
                :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  
361 VKLSGSVEVTQGVFLRVGAIAGTKGEFAIVSGKLEPGMTYTKLIDADVNNGINTSQVFIW 420  
                :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  
357 VTLGSKKV-T-GHVLSLPFGNKGNRRQEIEFGTLKPDTNYTSNEFSDSVDEGDLKVKFW 415  
                :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  
421 KGHLEFDSQNLGAEWINTSGKTYGKSTFCSDQIMCPNILQMLKPC 467  
                :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  
416' YNNVINLTLPKVGASK-ITVBENRGVSVPFCSETREDVLLILTAC 461  
                :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

T 12  
14

triacylglycerol lipase (EC 3.1.1.3) - guinea pig  
Species: Cavia porcellus (guinea pig)  
Date: 07-Oct-1994 #sequence\_revision 10-Nov-1995 #text\_change 18-Jun-1999  
Accession: S41084  
Lettinger, R., Thirstrup, K., Hjorth, S.; Boel, E.  
Journal: FEBS Lett. 338, 63-68, 1994  
Title: Cloning of the classical guinea pig pancreatic lipase and comparison with the I  
Reference number: S41084; MUID:94139896; PMID:8307159  
Session: S41084  
Molecule type: mRNA  
Accessions: 1-465 <CAR>  
References: EMBL:X77403; NID:G474916; PIDN:CAA54585.1; PID:G474917  
Description family: triacylglycerol lipase  
Keywords: carboxylic ester hydrolase

Protein Match      43.8%; Score 1108; DB 2; Length 465;  
Local Similarity    47.8%; Pred. No. 1.2e-81;  
Conservative         72; Mismatches 164; Indels 10; Gaps 9;

T 12

13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527  
 528  
 529  
 530  
 531  
 532  
 533  
 534  
 53

```

:ry Match      43.8%; Score 1108; DB 2; Length 465;
st Local Similarity 47.8%; Pred. No. 1.2e-81;
ches 225; Conservative 72; Mismatches 164; Indels 1

```

Qy	1	MLGIWIVAF	FFGTSRKEVCYERLGC	FDGLPWTFT	FTSTELVGLPWSPEK	INTREFL	Y	60
Db	1	MLLLMILS	-LLETVAGEVC	FDLGCFS	DNKFWAGTSR	PRKGLPWP	DSAINREFL	59
Qy	61	IENPNAYOEI	ISAVNSSTIOASY	FGTDKXIT	REINAGW	-KTDGKWORD	MNCVLOLED	118
Db	60	NENPNYORITA	-DSSVIRSD	FDKTRFTI	HGFI	DGEE	NWLADLCKAL	118
Qy	119	INLDWINGSRE	-YTHAVNNLR	VGAEVAYE	IDVLMKKFY	SPSKVHL	IGHSIGAH	177
Db	119	ICVDWREGSR	TLYSQASQNI	QVWGAEVAY	INFLQSLD	YPSSVHII	IGHSIGSHA	178
Qy	178	GSRIPG	-LGRITGLD	PAGPEHNT	KEVRLD	PSDANFV	DIHTNAAR	236
Db	179	GRTWGA	LRITGLD	PAEPFQY	TEI	VRILDP	SDAQFVD	238
Qy	237	GHLD	FYPNGG	XHMPGCE	DLITP	LKFNFNAY	KKEMASFF	296
Db	239	GHLD	FFPNGG	-QMPGCQ	KNILSQI	-VDIDG	WEGTRD	297
Qy	297	IAYPCRSY	TSFKA	NCNCF	CKSEGG	PTWGHF	ADRFHFK	356
Db	298	AGFSCDSY	SFSSNNK	PCPCAT	CPQMGHY	ADFPCKT	-KENFQ	356
Qy	357	HKLSYKLS	GSGBWTG	TGYFL	WVGAI	GKTB	EFAF	416
Db	357	YRIATLS	GSQV	-GHWVSL	FGDAGN	TQYEI	YRGS	415
Qy	417	QPIWKH	HLFED	SQNK	LAEW	INTSG	IKYKST	467
Db	416	KPIWNN	VINIT	TLPK	VASRIT	VT	RS	465

RESULT 13

triaclycerol lipase (EC 3.1.1.3) precursor, pancreatic - rabbit  
N:Alternate names: triglyceride lipase  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 31-Mar-2000  
C:Accession: JCI318

C/Accession: U0318  
R;Aleman-Gomez, J.A.; Colwell, N.S.; Sasser, T.; Kumar, V.B.  
Biochem. Biophys. Res. Commun. 188, 964-971, 1992  
A;Title: Molecular cloning and characterization of rabbit pancreatic triglyceride lipase  
A;Reference number: JCI318; MUID:93075235; PMID:1445366

A/Accession: U01318  
A/Molecule type: mRNA  
A/Residues: 1-465 <ALE>  
A/Cross-references: CB:M99365  
C/Superfamily: triacylglycerol lipase  
C/Keywords: carboxyl ester hydrolase  
F/1-15/Domain: signal sequence #status  
F/170\_194\_221/active site: Ser, Asp, His #status  
P:170\_194\_221/active site: Ser, Asp, His #status  
<SIG>

Query Match	43.8%;	Score 1107;	DB 1;	Length 465;
Best Local Similarity	48.1%;	Pred. No. 1.5e-81;		
Matches 230; Conservative	74;	Mismatches 150;	Indels 24;	Gaps 13;

Qy	1	MLGIWVAFJFFGTSRKVCYBERLGCFKDGLPWT-----RTFSTELVGLPWSPEKINTR	55
Db	1	MLLWALP-LILGAVAGLEYCYBERLGCFCNRIIPWSGGLTRPPST---LFTSPKDVNTR	55
Qy	56	FLLYTINPNAYQEISAVNSSTIQASYFGTDXTIRINIAGMTDCK--WORDMGNVLQL	113
Db	56	FLLYTNENPNFQISA--DASTIRGNGFTDRKTRPIIHGFTDKEENWLENLENFQV	114
Qy	114	EDINCINLDMINGSR-EYTHAVNLRVUGAEVAYFIDVLMKKFEPSPKVLHIGSLGAH	172
Db	115	ETVNCIVDWKGSRITYPQATQNIIRVGAEVAYLVGTQLSSLGVSPSNIHVIGSLGAH	174
Qy	173	LAGEAGSRIPG-LGRITGLDPAGPFFHNHTPKSVRLDPSDANFVDVHNTAARIIFELGVG	231
Db	175	AAGEVGRNTNGTIRITGLDPAPFPYFQGTPEVRLDPSDAQFVDVHTDAAPMVNLGFG	234



232 TIDACGLDFYPNGKHPGCGEDLITLLKFNFNAYKEWAFDCNHARSQFYAESIL 291  
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
235 MSQTGVGLHDFPFGKGEMPGCCKNVLSQI-VDINGVWEGRTRDFVACNLHRSYKYVADSIV 293  
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
292 NPDAFIAYPCRSYTEFPKAGNCFFCSKGCPTMGHFADRRHFKNMKTN--GSRYFLNTGSL 349  
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
294 NENGFAGFCASYTAFSANKCFPCS-NGCFQMGHYADRFP---SRKTDVGQTFYTLTGDS 349  
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
350 SPFAWRHLKSVLKSGSEVTGGTVFLRVGAIGKTEGFAIVSGKLEPGMTYTKLIDAVN 409  
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
350 SNFARWRVQAVTLSSRGVLT-GHYLVSLYSKGNKSKQYEFTGLLKPGDTHLNFEPSDVD 408  
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
410 VGNITSVQPIWKHLFEDSQNKLGAEWINTSGKYGYKSTFCSDQIMGNILQNLPKC 467  
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
409 VGDVQKVEFWNNVINFTLPKVASQ-IIVEQNDGRVFQFSTDIVREDILLITLTPC 465  
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

214

glycerol lipase (EC 3.1.1.3) - pig  
ies: Sus scrofa domestica (domestic pig)  
ies: 30-Nov-1980 #sequence revision 22-May-1981 #text\_change 31-Mar-2000  
ssion: A90638; A91256; A90675; A91124; A90732  
aro, J.; Boudouard, M.; Bonicel, J.; Guidoni, A.; Desnuelle, P.; Ravery, M.  
.m. Biophys. Acta 671, 129-138, 1981  
.e: Porcine pancreatic lipase. Completion of the primary structure.  
rence number: A90638; MUID:82113655; PMID:7326260  
ssion: A90638  
cule type: protein  
ues: 308-449 <DEC>  
xperimental source: pancreas  
ichetta, J.D.; Bidaud, J.; Guidoni, A.A.; Bonicel, J.J.; Ravery, M.  
f. Biochem. 97, 395-405, 1979  
.e: Porcine pancreatic lipase. Sequence of the first 234 amino acids of the peptide  
rence number: A91256; MUID:79236335; PMID:380992  
nts: carbohydrate-binding site  
ssion: A91256  
cule type: protein  
ues: 1-234 <AIA>  
xperimental source: pancreas  
loni, A.; Bonicel, J.; Bianchetti, J.; Ravery, M.  
mie 61, 841-845, 1979  
.e: Porcine pancreatic lipase. Sequence between the 235th and 307th amino acids.  
rence number: A90675; MUID:80088445; PMID:518929  
ssion: A90675  
cule type: protein  
ues: 235-307 <GUI>  
xperimental source: pancreas  
::: this sequence has since been revised at positions 302 and 305  
ouka, F.; Guidoni, A.A.; De Caro, J.D.; Bonicel, J.J.; Desnuelle, P.A.; Ravery, M.  
f. Biochem. 128, 331-341, 1982  
.e: Porcine pancreatic lipase. The disulfide bridges and the sulfhydryl groups.  
rence number: A91124; MUID:83105095; PMID:7151781  
nts: disulfide bonds  
ssion: A91124  
cule type: protein  
ues: 1-182, 184-449 <BEN>  
xperimental source: pancreas  
loni, A.; Benkroua, F.; De Caro, J.; Ravery, M.  
.m. Biophys. Acta 660, 148-150, 1981  
.e: Characterization of the serine reacting with diethyl p-nitrophenyl phosphate in  
rence number: A90634; MUID:82000578; PMID:6791692  
nts: annotation; substrate-binding site  
ent: Ser-152 reacts with emulsified or micellar diethyl p-nitrophenyl phosphate ar  
rfamily: triacylglycerol lipase  
ords: Carboxylic ester hydrolase; glycoprotein; lipid digestion  
1,237-261,285-296,299-304,433-449/Dissulfide bonds: #status experimental  
01/Dissulfide bonds: (or 90-103) #status experimental  
Binding site: carbohydrate (Asn) (covalent) #status experimental

Y Match 41.1%; Score 1039.5; DB 1; Length 449;  
Local Similarity 48.4%; Pred. No. 4.1e-76;

Matches	221; Conservative	62; Mismatches	157; Indels	17; Gaps	12;
Qy	19	EVCYERLGCFCQGLPWTRTFSTELVGLPWSPEK-INTRELLYTHNPAYOEISAVNSST	77		
Db	2	EVCFPRLGCFSDDAWAGIVQPLKILP--PDKVDVTRFLLYTNQNNYQELVA-DPST	58		
Qy	78	IQASYEGTKITRINIAGW--KTDGKQORDMCNVLLQLEDINCINLDWINGSRE-YIHAV	134		
Db	59	ITNSNFMQRKTRFIHGIDGEBEDJUSNICNLFKVESVNCICVDWKGSGRTGYTOAS	118		
Qy	135	NNLRVVVGAEVAYPIDVLMKKFYSPSKVHLIGHSLGAHLAGEAGSRIFG-LQRITGLDPA	193		
Db	119	QNIRIVGAEVAYFVEVLKSSILGYSPSNVHVIGHSLGSHAAGEACRRRTNGTIERITGLDPA	178		
Qy	194	GPFFHNTPEKVELDLSDSDANFVDVITNAAIRILPELGVGTIDACGHLDPYNGGKMPQCE	253		
Db	179	EPFCQGTPELVELDLSDSDAKFVDVITNDAPIINPLGFGMSQTVGHLDFFPFGGKMPQCG	238		
Qy	254	DLITPLLLKFNFNAYKEMASFPDCNHARSYOFYAESILNPDAFIAYPCRSYTSFKAGNCF	313		
Db	239	KNILSQI-VDIDIGIEGTRDFVACNHLRSYKYADSIILNPDGAGFPDSSYNVTANKCF	297		
Qy	314	PCSKECPTMGHFADRFHQWKTNG--SHVFLNTGSLSPFARWRHKLSVKLSSEVTQG	371		
Db	298	PCSECCPQMGHYADRF---PGKTNGVQVFLYNTGDSANFARMRYKYSVTLGSKKVT-G	353		
Qy	372	TVFLRVGGAIGTKGFPAIVSGKLEPGMTYTKLIDADVNVGNITSVOFTW-KKHLFEDSCN	430		
Db	354	HIILVSLFNGEGRQVEIYKGTLPQDNTSHDEFFSDVEVGLQKVFLWYNNVINPTLP	413		
Qy	431	KLGAEMVINTSGYKVKSTFCQSDIMGNNILCNLKPC	467		
Db	414	RVGASK-IIVERNDGKVFDFCSQETVREEVLLTNPC	449		

RESULT 15

A28997 triacylglycerol lipase (EC 3.1.1.3) precursor, hepatic - human  
N:Alternate names: hepatic triglyceride lipase; lipase; triacylglycerol acylhydrolase, he  
C:Species: Homo sapiens (man)  
C>Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text\_change 18-Jun-1999  
C:Accession: A28997, S00201; A33553; A29932; A35740  
R:Martin, G.A.; Busch, S.O.; Meredith, G.D.; Cardin, A.D.; Blankenship, D.T.; Mao, S.J.T.  
on, R.L.  
J. Biol. Chem. 263, 10907-10914, 1988  
A:Title: Isolation and cDNA sequence of human postheparin plasma hepatic triglyceride li  
A:Reference number: A28997; MUID:88273217; PMID:2839510  
A:Accession: A28997  
A:Molecule type: mRNA  
A:Residues: 1-499 <VAR>  
A:Cross-references: GB:J03895; NID:G339594; PIDN:AAA61165.1; PID:G339595  
R:Stahnke, G.; Sprenger, R.; Augustin, J.; Will, H.  
Differentiation 35, 45-52, 1987  
A:Title: Human hepatic triglyceride lipase: cDNA cloning, amino acid sequence and expres  
A:Reference number: S00201; MUID:88112595; PMID:2828141  
A:Accession: S00201  
A:Molecule type: mRNA  
A:Residues: 1-255, F', 257-499 <STA>  
A:Cross-references: EMBL:X07228; NID:G32497; PIDN:CAA30188.1; PID:G32498  
R:Cai, S.O.; Wong, D.M.; Chen, S.H.; Chan, L.  
Biochemistry 28, 8966-8971, 1989  
A:Title: Structure of the human hepatic triglyceride lipase gene.  
A:Reference number: A33553; MUID:90105435; PMID:2605236  
A:Accession: A33553  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 30-255, F', 257-499 <CAI>  
A:Cross-references: GB:M29187; GB:J02882; GB:M29188; GB:M29189; GB:M29190; GB:M29191; GB:  
R:Patra, S.; Luo, C.C.; Li, W.H.; Vantunen, P.; Leebetter, D.H.; Brown, M.A.; Chen, S.H.  
J. Biol. Chem. 263, 1107-1110, 1988  
A:Title: Human hepatic lipase. Cloned cDNA sequence, restriction fragment length polymor

```

Job completed: January 6, 2004, 11:23:40
Time : 31 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

rotein - protein search, using sw model

1: January 6, 2004, 10:17:33 ; Search time 60 Seconds  
(without alignments)  
2008.509 Million cell updates/sec

US-10-038-517-2

Score: 2527

nce: 1 MLGIWIAFLFFGRSGKEV.....STFCSQDIMPNTLQNLKPC 467

ng table: BLOSUM62

Gapop 10.0 , Gapext 0.5

hed: 830525 seqs, 258052604 residues

number of hits satisfying chosen parameters: 830525

am DB seq length: 0

am DB seq length: 2000000000

processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

ase : SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Score	Query Match	Length	DB ID	Description
1180	46.7	473	11 O70478	O70478 mus musculus
1162	46.0	467	4 O8TAU2	O8tau2 homo sapien
1121	44.4	434	11 P81139	P81139 cavia porce
1108	43.8	470	13 O8QGM1	O8qgm1 anguilla ja
1106	43.8	465	11 Q9QWF3	Q9qwf3 spermophilu
1098	43.5	465	11 O88354	O88354 spermophilu
1084	42.9	452	6 Q95KP4	Q95kp4 equus cabal
1063	42.1	427	6 Q28287	Q28287 canis famil
1062	42.0	451	11 Q9QUN4	Q9qun4 rattus norv
553.5	22.0	952	5 Q9VX01	Q9vxo1 drosophila
513	20.3	451	4 Q8WYX8	Q8wyy8 homo sapien
503	19.9	510	11 Q8VC44	Q8vc44 mus musculu
501.5	19.8	499	6 Q46559	Q46559 oryctolagus
500	19.8	540	5 Q9W448	Q9w448 drosophila
478.5	18.9	452	6 Q9BDJ4	Q9bdj4 oryctolagus
476.5	18.9	510	13 Q98UI2	Q98ui2 pagrus majo

17	476.5	18.9	511	13	Q98UI3	Q98ui3 pagrus majo
18	475.5	18.8	451	11	O8CIV3	O8civ3 mus musculu
19	473.5	18.7	449	11	O8BXB5	O8bxb5 mus musculu
20	460.5	18.2	490	13	Q9PSV4	Q9psv4 gallus gall
21	452	17.9	475	4	Q96FC4	Q96fc4 homo sapien
22	449	17.8	449	6	Q95KN9	Q95kn9 macaca fasc
23	445	17.6	503	13	Q9MEY2	Q9mey2 oncorhynch
24	442.5	17.5	474	11	O8C562	O8c562 mus musculu
25	403.5	16.0	456	11	P97535	P97535 rattus norv
26	402.5	15.9	376	4	Q9UPD2	Q9upd2 homo sapien
27	402.5	15.9	455	4	O95991	O95991 homo sapien
28	401.5	15.9	323	4	O8TBC7	O8tec7 homo sapien
29	398.5	15.8	456	11	Q99U51	Q99u51 mus musculu
30	397.5	15.7	456	11	O8VI78	O8vi78 mus musculu
31	385.5	15.3	421	11	O8CI45	O8ci45 mus musculu
32	380.5	15.1	332	6	O77644	O77644 pan troglod
33	373	14.8	484	5	Q9VJH1	Q9vjh1 drosophila
34	369.5	14.6	411	5	O9VLU1	O9vlu1 drosophila
35	360.5	14.3	163	11	Q9D950	Q9d950 mus musculu
36	359.5	14.2	394	5	Q9V7V4	Q9v7v4 drosophila
37	358.5	14.2	390	5	Q9V9P0	Q9v9p0 drosophila
38	341	13.5	675	5	Q9VB94	Q9vb94 drosophila
39	339	13.4	575	5	Q9VOC3	Q9voc3 drosophila
40	336.5	13.3	337	5	Q9VB91	Q9vb91 drosophila
41	325.5	12.9	404	5	Q9W304	Q9w304 drosophila
42	324.5	12.8	341	5	Q9VB89	Q9vb89 drosophila
43	324.5	12.8	404	5	Q8MYX6	Q8myx6 drosophila
44	323	12.8	337	5	Q9VB92	Q9vb92 drosophila
45	318.5	12.6	172	6	Q9TSB3	Q9tsb3 ovis aries

#### ALIGNMENTS

#### RESULT 1

O70478 ID O70478 PRELIMINARY; PRT; 473 AA.  
AC O70478;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Pancreatic lipase related protein 1.  
GN PNLIPLP1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Swiss Webster; TISSUE=Exorbital lacrimal gland;  
RX MEDLINE=99249620; PubMed=10235541;  
RA Remington S.G., Lima P.H., Nelson J.D.;  
ET "Pancreatic lipase-related protein 1 mRNA in female mouse lacrimal gland.";  
RL Invest. Ophthalmol. Vis. Sci. 40:1081-1090(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Pancreas;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
DR EMBL; AF061274; AAC15774.1; -  
DR EMBL; AK028105; BAC25750.1; -  
DR HSSP; P06857; IRP1.  
DR MGD; MGI:197723; Pnliplr1.  
DR InterPro; IPR000734; Lipase.  
DR InterPro; IPR001024; Lipoxigenase LH2.  
DR InterPro; IPR000379; Ser esters\_site.  
DR Pfam; PF00151; lipase; 1.  
DR Pfam; PF01477; PLAT; 1.

PRINTS; PR00821; TAGLIPASE.  
SMART; SM00308; LH2; 1.  
PROSITE; PS00120; LIPASE\_SER; 1.  
SEQUENCE 473 AA; 52695 MW; PB11C08E6BC2763 CRC64;

!y Match  
t Local Similarity 46.7%; Score 1180; DB 11; Length 473;  
ches 231; Conservative 73; Mismatches 159; Indels 8; Gaps 6;

1 MLGIWIVAF.FFGTSRKVECYERLCGPKDGLPWRTRFSTELVGLPSPEKINTREFLYT 60  
1 MLILWTIPLFLGAAQCKEVCYDNLGCSDAEPWAGTAIRPLKLLPSPKINTREFLYT 50

61 IHNPNAYQETISAVNSSTICASTPGTDKITRINIAGW--KTGGKWQRDMCNVLQLEDINC 118  
61 NENPTAFTQLQLSDPSTIEASNFQVARKTRFIHGFIDKGEEVNVDMCKNMFQVEVNC 120

119 INLDWNGSR-EYIHAVNNLRVVUGAEVAFIDVMKKFEYSPSKVHLIGHSLGAHLAGEA 177  
121 ICVDWKRGSGTTITQAANNVRVVGVAQMIDILVRNPENTSAKVHLIGHSLGAHVAGEA 180

178 GSRIPLGLGRITGDPAQPFPHNTPEVKRLDPDSDFANFDVVIHTNAARILFELGVGTIDACG 237  
181 GSRTFGLGRITGDDPYEANFEGTPSEVRLDPDSADFVDVIHTDAAPLIPLFGFGTNMVG 240

238 HLDVPYNGGXHMFGCC-EDLITPLLKFENNAYKEMKMAFFDCNHARSQVFAESILNDPAF 295  
241 HFDFPNGGYMGCGKNALSQIV--DIGTWGSTRDFVACHLRSIKYLESIILNPGF 298

297 IAYPCRYSTFSKAGNCFPPCKSGCGCTMGHFADRFPHFKMKMTKNGSHYFLNTGSLSPFARWR 356  
299 AAYPCASYDRFESNKFCPPCDQGCPQMGHYADKF-ANNTSVSEPOKFFLTNGAKXFNARWR 357

357 HKLSVKLSGSSEVTQGVTVFLRVGGA.GTKGEFAIVSGKLEPGMYTYKLIDADVNGVNITSV 416  
358 YRLVSTFSGRTVT-GQVKVSLFGSNGNRQCDIFRGIKFGATHNEFDAKDVGDTIEKV 416

417 QFIWKXHLFSDSNKLGAEMWINTSGKYGKYKSTFCSDQIMGNILQNLPXC 467  
417 KFLNNHVVNFPSPKVGAAKITVQKGEERTEHNFCSEETVREJILLTLPLCP 467

T 2  
/2

PRELIMINARY; PRT; 467 AA.

Q8TAU2  
Q8TAU2;  
01-JUN-2002 (TrEMBLrel. 21, Created)  
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

Pancreatic lipase-related protein 1.  
Homocaptens (Human).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
NCBI\_taxid=9606;  
[1]  
SEQUENCE FROM N.A.  
TSSB=pancreas;  
Straussberg R.;  
Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
EMBL; BC025784; AAH25784.1; -  
InterPro; IPRO00734; Lipase.  
InterPro; IPRO01024; Lipoxigenase\_LH2.  
InterPro; IPRO00379; Ser\_estrs\_site.  
Pfam; PF00151; lipase; 1.  
Pfam; PF01477; PLAT; 1.  
SMART; SK00308; LH2; 1.  
PROSITE; PS00120; LIPASE\_SER; 1.  
SEQUENCE 467 AA; 51833 MW; 7680EA1E22E258A3 CRC64;

y March  
t Local Similarity 46.0%; Score 1162; DB 4; Length 467;  
ches 229; Conservative 75; Mismatches 159; Indels 8; Gaps 6;

Qy	1	MLGIWIVAFLEFFGTSRKEVCYERLGCDFKGLPWTRTFTSELVGLPWSPEKINTRELLYT	60
Dd	:	: :	
Dd	1	MLIPFWTITFLGAAGKEVCYEDLGCFSDTEPWGCTAIRPLKILPWSPEKIGTRFLLYT	60
Qy	61	IHPNAYOEISAVNSSTIQASYGTDKITRINTAGM--KTDGKWQRDMCNVLLOLEDINC	118
Dd	61	NENPNNFQIILLSDPSTIEASNFQMDRKTRFIHGFDIDKGDSWVTDMCKLPFEVEVNC	120
Qy	119	INDLWINGSR-EYIHAVNNLRVVGAEVAYFIDLVMKKKFYSPSKVHLIGHSLGAHLAGEA	177
Dd	121	ICVDWKKSQAITYQANNRVVGAOVAQWLDDLTLLEYSPYPSPKVHLIGHSLGAHVAGEA	180
Qy	178	GSRIPGLGRITGLDPAGPFPHNTPKPEVRLDPSDANFVDVIHTNAARILFELVGTTIDACG	237
Dd	181	GSTPGLSRITGLDPEVASPESTEEVERLDPSDAFVDVIHTDAAPLIPLFGFTNQMG	240
Qy	238	HLDYPNGGHMPGC-EDLTPLLKNFNAYKEMASFPDCNHARSQVFAESILNPDAF	296
Dd	241	HLDFPNGSGMEFCCKNALSQIV--DLGINAGTRDFVACHNLRSYKYLESILNPDG	298
Qy	297	IAYPCRSYTSFKAGNCCFFCSKECPTGMHFAFRPHFKNMKTNGSHYFLNTGSLSPFARWR	356
Dd	299	AAYPC"SYXSFSKDCCFPDQCPOGWYADKFACTRGSBEQ-QKFFLNTGEASNFAWR	357
Qy	357	HKLSVKLSGEVTVGGTFVLEVGAIGCTEPAIVSKLEPGMYTKLIPADNVNGITSV	416
Dd	358	YGVSITLSGRAT-GQIKVALFNKGNTHQYSIFRGILKPSTHSYEPDAKLDVGDTDKV	416
Qy	417	QFTWKHLFEDSONKLGAEWINTSGYKYSTFCSDIMGNILNKXPC	467
Dd	:	: :	
Dd	417	KFLWNVNINPTLPKVGATKITVQKGEKTVYNFCSDDTVREDTLTLTPC	467

RESULT 3

ID	P81139	PRELIMINARY;	PRT; 434 AA.
AC	P81139;		
DT	01-JAN-1998 (T+EMBLrel. 05, Created)		
DT	01-JAN-1998 (T+EMBLrel. 05, Last sequence update)		
DT	01-MAR-2002 (T+EMBLrel. 20, Last annotation update)		
DE	Pancreatic lipase related protein 2 (SC 3.1.1.3) (GPL) (Fragment).		
OC	Cavia porcellus (Guinea pig).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Myricognathi; Caviidae; Cavia.		
RN	NCBI_TaxId=10141;		
RP	[1]		
RC	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.		
RA	TISSUE=PANCREAS;		
RX	MEDLINE=93257433; PubMed=8490016;		
RR	Hjorth A., Carriere F., Cudrey C., Woldike H., Boel E., Lawson D.M.,		
RT	Ferrato F., Cambillau C., Dodson G.G., Thim L., Verger R.;		
RL	"A structural domain (the lid) found in pancreatic lipases is absent		
RL	in the guinea pig (phospholipase.";		
RT	Biochemistry 33:4702-4707(1993).		
CC	-1- CATALYTIC ACTIVITY: TRIACYLGLYCEROL + H(2)O = DIACYLGLYCEROL + A		
CC	FATTY ACID ANION.		
CC	-1- TISSUE SPECIFICITY: PANCREAS.		
CC	-1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,		
CC	HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).		
DR	HSP; P24318; 1BUB.		
DR	InterPro; IPR000734; Lipase.		
DR	InterPro; IPR001024; Lipoxigenase_LH2.		
DR	InterPro; IPR000379; Ser_estr_site.		
DR	Pfam; PF00151; lipase; 1.		
DR	Pfam; PF01477; PLAT; 1.		
DR	PRINTS; PR00821; TAGLIPASE.		
DR	SMART; SM00308; LH2; 1.		
DR	PROSITE; PS00120; LIPASE_SER; 1.		
KW	Hydrolase; Lipid degradation; Pancreas; Glycoprotein.		
FT	NON TER		
FT	ACT_SITE 154 154		
FT	ACT_SITE 178 178		
FT	ACT_SITE 247 247		
FT	CHARGE RELAY SYSTEM (BY SIMILARITY).		
FT	CHARGE RELAY SYSTEM (BY SIMILARITY).		
FT	CHARGE RELAY SYSTEM (BY SIMILARITY).		

DR SMART; SMO0308; LH2; 1.  
DR PROSITE; PS00120; LIPASE\_SER; 1.  
KW Hydrolase.

SQ SEQUENCE 470 AA; 52676 MW; 2F32DC4A65B0F3F3 CRC64;

Query Match 43.8%; Score 1108; DB 13; Length 470;  
Best Local Similarity 45.0%; Pred. No. 3.4e-87;  
Matches 213; Conservative 78; Mismatches 166; Indels 16; Gaps 8;

QY 4 IWIIVAFLEFGTSGRKEVCYERLGCDFKGLPWTFTSTELVCLPWSPKINTRFLLYTIHN 63  
:  
Db 5 VWILGLEFLGTASGAECVENIGCFTDDVPWAGTTERPIAKLPWSPEKIGTRFWLPTQN 64

QY 64 PNAVOEISAVNSSITQASYFGTDKITRINTAGW--KTGKWQRDMCNVLLOLEDINCINL 121  
:  
Db 65 PNNHOEITTKEDIIILASNYNGTRK-TREITHGVYVDKGDNWLMDCKMLQLQVEDINCICV 123

QY 122 DWINGSRE-YTHAVNRLRVGAEVAYIDVLMKKFEYSPSKVHLIGHSLGAHLAGSAGR 180  
:  
Db 124 DWKGGRLTYQSASNIRVIGAQMAYMIOLFQKYVQRPESVHIIGHSLGAHCAGEARR 183

QY 181 IPGLGRITGLDPAGPFPHNTPKVEVRLDPSDANFVDVIHTNAARILPELVGTTIDACGHL 240  
:  
Db 184 TPNLGRIITGLDPABPYQGCPSLVRLDPSDAKFVDVHTDKAPMIPLYLGMAQAAGVCHLD 243

QY 241 FYPNGSKHPGCE-DLTPLLKFNFNAYKKEMASFFDCNHARSQVFAESILNPDAFIAY 299  
:  
Db 244 FYPNGGEMPCDKXNILSQTV--DIDGIWEGTRDFVACNELRSYKIYSDSILNPEGTGY 301

QY 300 FORSYTSFKAGNCFCSCKEGCTMGHFADRFHKNMKTNGSH---YFLNTGSLSPFARWR 356  
:  
Db 302 PCSDGVFSRSCRCPCDGAGCPFMGHADKFR---RPGAENKVKFYLNADAKPFGRYR 357

QY 357 HKLSVKLSGSB--VTQGTVFLRVCGAIKGTGEPAIVSGKLPERGMTYTKLIDADVNGNIT 414  
:  
Db 358 YKVVTVTIRADALLTGTMSVAIYGTQGNTRQYRKGLLKPGNTVEAYIDTETDAGEVT 417

QY 415 SVQFTWKHLPEDSQNKLGAEWINTSGKYGKYSTQCSODIMGPNILQNLKPC 467  
:  
Db 418 KMKFTWDNSVINWPFKLGAEKIVLRQGDRRIFSCSSSTVEEMLQTVPCC 470

RESULT 5  
Q9QWF3 PRELIMINARY; PRN; 465 AA.

ID Q9QWF3 AC Q9QWF3  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Pancreatic triacylglyceride lipase.  
GN PTL.  
OS *Spermophilus tridecemlineatus* (Thirteen-lined ground squirrel).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;  
OC *Spermophilus*.  
OX NCBI\_TaxID=43179;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=White adipose tissue;  
RA Bauer V.W., Andrews M.T.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF177402; AAD51123.2; -.  
DR HSSP; FI6233; LLPB.  
DR InterPro; IPR000734; Lipase.  
DR InterPro; IPR001024; Lipoxigenase\_LH2.  
DR InterPro; IPR000379; Ser\_estrs\_site.  
DR Pfam; PF00151; lipase; 1.  
DR Pfam; PF01477; PLAT; 1.  
DR PRINTS; PR00821; TAGLIIPASE.  
DR SMART; SMO0308; LH2; 1.  
DR PROSITE; PS00120; LIPASE\_SER; 1.  
SQ SEQUENCE 465 AA; 51235 MW; 4A1F0043E6F54A15 CRC64;



SEQUENCE FROM N.A.  
 "Jayne S., Kerfelec B., Foglizzo E., Chapus C., Crenon I.;  
 "Horse pancreatic lipase related protein 2 : expression in adult and  
 properties.";  
 Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 EMBL; AF359386; AAK38605.1; -;  
 InterPro; IPR000734; Lipase.  
 InterPro; IPR001024; Lipoxigenase LH2.  
 InterPro; IPR000379; Ser\_estrs\_site.  
 Pfam; PF00151; lipase; 1.  
 Pfam; PF01477; PLAT; 1.  
 PRINTS; PR00821; TAGLIPASE.  
 SMART; SM00308; LH2; 1.  
 PROSITE; PS00120; LIPASE\_SER; 1.  
 VONTER  
 SEQUENCE 452 AA; 5063 MW; 386AFA5732D69AD9 CRC64;  
 y Match 42.9%; Score 1084; DB 6; Length 452;  
 : Local Similarity 44.6%; Pred. No. 3.7e-85;  
 :hes 209; Conservative 80; Mismatches 144; Indels 36; Gaps 8;  
 18 KEVCYERLCGKFDGLPWTPTSTSTELVGLPWSPEKINTRELLYTHNPNAYQEISAVNSST 77  
 1 KEVCYTPGCGSDDKPMAGTLQRLPLKSLPWSPEEVTNTRFLYTNKNDPSYQLITARDVAT 60  
 78 IQASVFGTDKTRINIAIGKTDGK--WORDMCNVLQLEDINCINLNDWINGSR-EYIHAV 134  
 51 IKSSNFQSRKTHFVHGFDRGDSWPSDMCKILQVETTCISVDWSSGAKAEYQAV 120  
 135 NNLRVVGAEAFFIDVLMKKFEYSPSKVHLGHSGLAGHAGEAGSRIPG-LGRITGLDPA 193  
 121 QNIRIVGAETAYLQILLTELSPENPVNHHGSLGAHTAGEAGRLRGVRVTGLDPA 180  
 194 GPFFENTPEKVRLPDSDANFVDVHTHNAARILFELGVTGTTDACHLDLPNGSKHPGCE 253  
 181 ECFQDAEEVRLPDSDAQFVDVHTDASPLPGLSGMGSQKQVGHMDFPENGKQMPGCK 240  
 254 DLITPLKLFNFNAYKEMASFFD-----CNHARSQFYAESILNPDFAIYAP 300  
 241 -----RSSFSFTIDINGIWQGAQDYLACNHLKSEYYSILNPDGFLAYP 286  
 301 CSYTSFYAGNCPCSKGEGCTMGHFADRPHFKMKTNGSHYFLNTGSLSPFARMHKLK 360  
 287 CDSYDKFOENGCFPCPAGGCPMGHYADQYKEKTSAVEQT--FFLNTGESGDYTSWRYVS 345  
 361 VKLSGSEVTQGTVFLRVGGAIGKTGEFAIVSGKLEPGMTYTKLIDADVN--VGNITSVOF 418  
 346 ILAGSGKANGLYKTVLRGSGNGSKQYEIFKGLSPDSSTI--LDVDVNFIGKIOEVKF 403  
 419 IWKXHLFEDSQNKLGAEMVINTSGYKGYSTFCSDIMGNPILQNLKPC 467  
 404 VNNKTVLNLKLPQLGASRITVQSGADGTEYKFGSGTVQDNVQSLVPC 452  
 ; 8  
 228287 PRELIMINARY; PRT; 427 AA.  
 228287;  
 01-NOV-1996 (TRENBLrel. 01, Created)  
 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 pancreatic lipase.  
 :anis familiaris (Dog).  
 :ukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 :mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 :CBI\_TaxID=9615;  
 :SEQUENCE FROM N.A.  
 :EDUINP=89327249; PubMed=2502543;  
 :Hickel F.S., Weidenbach F., Swarovsky B., LaForge K.S., Scheele G.A.;  
 :Structure of the canine pancreatic lipase gene.";  
 :. Biol. Chem. 264:12895-12901(1989).  
 :MBL; M28151; AAA30840.1; -;

DR EMBL; M28141; AAA30840.1; JOINED.  
 DR EMBL; M28142; AAA30840.1; JOINED.  
 DR EMBL; M28143; AAA30840.1; JOINED.  
 DR EMBL; M28144; AAA30840.1; JOINED.  
 DR EMBL; M28145; AAA30840.1; JOINED.  
 DR EMBL; M28146; AAA30840.1; JOINED.  
 DR EMBL; M28147; AAA30840.1; JOINED.  
 DR EMBL; M28148; AAA30840.1; JOINED.  
 DR EMBL; M28149; AAA30840.1; JOINED.  
 DR EMBL; M28150; AAA30840.1; JOINED.  
 DR HSSP; P06857; 1RP1.  
 DR InterPro; IPR000734; Lipase.  
 DR InterPro; IPR001024; Lipoxigenase LH2.  
 DR InterPro; IPR000379; Ser\_estrs\_site.  
 DR Pfam; PF00151; lipase; 1.  
 DR Pfam; PF01477; PLAT; 1.  
 DR PRINTS; PR00821; TAGLIPASE.  
 DR SMART; SM00308; LH2; 1.  
 DR PROSITE; PS00120; LIPASE\_SER; 1.  
 SQ SEQUENCE 427 AA; 47145 MW; 93F3CD4131C15CED CRC64;  
 Query Match 42.1%; Score 1063; DB 6; Length 427;  
 Best Local Similarity 44.9%; Pred. No. 2.2e-83;  
 Matches 211; Conservative 68; Mismatches 145; Indels 46; Gaps 5;  
 QY 1 MLGTVIYAFLEPFGTSRCKEYERLCGKFDGLPWTPTSTSTELVGLPWSPEKINTRELLYT 60  
 DB 1 MVSITWTFALFLGAAKAEVCYEQICGFSDAEFWAGTAIRPLKVLFWSPER-GTRFLLYT 60  
 QY 61 IHNFNAYQEISAVNSSTIQASYFGTQDKITRINIAIGW--KTDGKQWQDMCNVLQLEDINC 118  
 DB 61 NKNENNFQTLPSDPSTIEASNFQDKTETFIHGFIDKGEENWLLDMCKNFKVEEYVC 120  
 QY 119 INLDWINGSR-EYIHVNNLRVVGAEVAFIDVLMKKFEYSPSKVHLGHSGLAGHAGEA 177  
 DB 121 ICVDKKGSGTSYTOAANNRVVVGAAQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 180  
 QY 178 GSRTPGLGRITGLDPAGPFPHTPKVRLDPDSDANFVDVHTHNAARILFELGVTGTTDAG 237  
 DB 181 GSRTPGLGRITGLDPVEASFQGPTEVRDLPDADFDVHTHNAARILFELGVTGTTDAG 233  
 QY 238 HLDYFPGKGMPCGEDLITPLKFNFNAYKEMASFFDCNARSQFYAESILNPDFAI 297  
 DB 234 -----RDFVACNELRSYKYSIESILNPDGFA 259  
 QY 298 AYPGRSYTSFKAGNCPCSKGEGCTMGHFADRPHFKMKTNGSHYFLNTGSLSPFARMH 357  
 DB 260 SYPCASYRAPESNKKPCPCPDQCPQMGHYADKFAVKT-SDETQKFLNTGDSNFAWR 318  
 QY 358 KLSVKLSGSEVTQGTVFLRVGGAIGKTGEFAIVSGKLEPGMTYTKLIDADVNNGNITSVQ 417  
 DB 319 GVSTILSGKEAT-GQAKVALFGSKGNTHQFNIFKGLKPGSTHSNEFFDAKLDVGTIERK 377  
 QY 418 FTKXHLFEDSQNKLGAEMVINTSGYKGYSTFCSDIMGNPILQNLKPC 467  
 DB 378 FLWNNVNVNPTFPKVGAAKITVQKBEKTVHFCSESTVREDVLLTLTPC 427  
 RESULT 9  
 Q9QUN4  
 ID Q9QUN4 PRELIMINARY; PRT; 451 AA.  
 AC Q9QUN4  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE Pancreatic lipase (Fragment).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WISTER-KING; TISSUE=Pancreas;



Tsujita T., Sumida M., Sumiyoshi M., Kameda K., Okuda H.;  
 "Direct evidence that alkaline lipase from Rat brain is the same  
 enzyme as pancreatic lipase from Rat Pancreas."  
 Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.

[2]  
 SEQUENCE FROM N.A.  
 STRAIN=WISTER-KING;  
 Sumida M.;  
 Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.

[3]  
 SEQUENCE FROM N.A.  
 STRAIN=WISTER-KING;  
 Takahiro T.;

"Direct evidence that alkaline lipase from Rat brain is the same  
 enzyme as pancreatic lipase from Rat Pancreas."  
 Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 EMBL; D88535; BAA13638.1; -  
 EMBL; D88534; BAA13637.1; -  
 HSP; P16233; 1LBP.

InterPro; IP000734; Lipase.  
 InterPro; IP001024; Lipoxigenase LH2.  
 InterPro; IP000379; Ser\_estrs\_site.  
 Pfam; PF00151; lipase; 1.  
 Pfam; PF01477; PLAT; 1.  
 PRINTS; PR00821; TAGLIPASE.  
 SMART; SM00308; LH2; 1.  
 PROSITE; PS00120; LIPASE\_SER; 1.  
 NON TER 451 451  
 SEQUENCE 451 AA; 49939 MW; 65D19706B5D5B732 CRC64;

Query Match 42.0%; Score 1062; DB 11; Length 451;  
 t Local Similarity 47.0%; Pred. No. 3e-83;  
 ches 215; Conservative 74; Mismatches 158; Indels 10; Gaps 9;

1 M L G I W I V A F F G T S G K E V C R I G C F K D G L P W T R T S T E L V G L P W S P E K I N T R F L Y T 60

1 M L M L W T F A V L - L G A V A G E V C F K L G C F S D D A P W S G T I D R P L K A L P W S P A Q I N T R F L Y T 59

61 I H N P N A Y Q B I S A N S T I O A S V F G T D K I T R I N A G W - - K T D G K W O R D M C N V L L O E D I N C 118

60 N E N Q D Y Q K I T S - D A S I R N S F K T N K R T I I I H F I D K G E N W L S D M C K N F V S V N C 118

119 I N L D W I N G R - E Y I H A V N N L R V G A E V A F I D V L M K F E Y S P K V H L I G H S L G A H L A G E A 177

119 I C V D W K G S R A T Y T Q A T O N V R V V G A E V A L L N V L K S D L G S P D N V H L I G H S L G S H V A G E A 178

178 G S R I P G - I G S I T L D P A G F F E H T K E V E L D S D A N F V D V I H T N A R I L P F L G V G T T D A C 236

179 G K T F G A I G R I T L D A E Y F O G T P E E V E L D P T D A Q F V D A I H T D A P I I P N L G F M S Q T V 238

237 G H L D F P N G S K G M P G C E D L I T L L K F N F N A Y K E M A S F F D C N H A R S Y Q F Y A E S I L N P D A F 296

239 G H L D F P N G M E M P G C Q K N I L S Q I - V D I D G I W E G T R D F A A C N H L R S Y K Y Y T D S I V N P T G F 297

297 I A P C R S Y S F R A G N C F F S K G C C T M G H A D R F F K W K N T G N G H Y F I N T G S L G F P A R W R 356

298 S G F S C S S Y N F S A N K C F P C G S G C C P Q M G H A D K Y P G K T - K E L Y Q K F Y L N T D K G N F A R W R 356

357 H K L S V K L S G S E V T Q G T V F L R V G G A I G K T G E F A I V S G K L E P G M T V T K L I D A D V N G N I T S V 416

357 Q Y V T V I L S G Q V T - G H I L V S L F G N G S N K Q E V F K S L H P G D T H V K F D S D M D V D L Q K V 415

417 Q F T W K G L F D S Q N K L G A E W I N T S G K Y G K S T F C S Q 453

416 K F I W Y N V I N P L P K V G A S R - I S V E R N D R V F N F C S Q 451

T 10

Q3VX01

Q3VX01;

01-MAY-2000 (T-EMBLrel. 13, Created)

01-MAY-2000 (T-EMBLrel. 13, Last sequence update)

DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)  
 DE C6847 protein.  
 GN C6847

OS Drosophila melanogaster (Fruit fly).

OC Rukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OC NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RK MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton R.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M.R., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimball B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).

DR EMBL; AE003507; AAF48784.1; -

DR HSP; P29183; 1HPL.

DR Flybase; FBgn0030884; C6847.

DR InterPro; IP000734; Lipase.

DR InterPro; IP000379; Ser\_estrs\_site.

DR Pfam; PF00151; lipase; 1.

DR PRINTS; PR00821; TAGLIPASE.

SQ SEQUENCE 952 AA; 104865 MW; EAC08569B95455CD CRC64;

Query Match 22.0%; Score 555.5; DB 5; Length 952;

Best Local Similarity 31.4%; Pred. No. 5.9e-39;

Matches 148; Conservative 72; Mismatches 156; Indels 95; Gaps 14;

QY 18 KEVCYERLCGFCGDLGFLPWTFTSTELVGLPWSPEKINTRFLLYT-----IIN 63

Db 22 KVCYVGLGCFEDSGPF-----AYLEMLPSSPEINTKFFYSTRQSRDRLMELFLNM 76

QY 64 PNAVQ-----EISAVNSSTIOA-----SY 82

Db 77 TNAPFRGRETEVSTSSPEGTSGRSSVASAPSSAVNATFTTTERFGGQKFTPSIDLL 136

QY 83 FGTDKIT-RINIAWKT--DGKWORDCNVLLDINCINLDWINGSR--EYIHAVNVL 137

[illegible]

291 LNPD-AFIATPCRSYTSFKAGNCFPSKGCPTMGHFAHFRHFKNMKTNGSHYFLNTGSL 349  
 292 QHSLSQSIGFQCSDMGFSQCLCLSKCKGRCNTLGYDIR---XDRSGSKSLFLITRAQ 347  
 350 SPFAWRHKLKSVKLSG--SEVTQGVFLRVGGAIGKGTGFAIVSGK-LPFGMTYTKLIDA 406  
 348 SPFKVHYQKFIQFINGIEKRPVPTTMSLLGTKEIKRIPITLGEIGTSNKYSFLITL 407  
 407 DVAVNGNITSVQFIWK--HLFEDSQNKL--GAE-----MVINT-----SGKYGYKSTF 450  
 408 KDTIGELILLKFKWNSAVWVWNTVQITLWGIPEHSGILKLTINWVKGAGEQRTQRTF 467  
 451 CSQDI 455  
 468 CPENL 472

PRELIMINARY; PRT; 499 AA.

01-JUN-1998 (TrEMBLrel. 06, Created)  
 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 Hepatic lipase.  
 Oryctolagus cuniculus (Rabbit).  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 NCBI\_TaxID=9986;  
 [1]  
 SEQUENCE OF 3-499 FROM N.A.  
 MEDLINE=92121762; PubMed=1770315;  
 Warren R.J., Ebert D.L., Mitchell A., Barter P.J.;  
 "Rabbit hepatic lipase cDNA sequence: low activity is associated with  
 low messenger RNA levels";  
 J. Lipid Res. 32:1333-1339(1991).  
 [2]  
 SEQUENCE OF 1-2 FROM N.A.  
 Mitchell A.;  
 "First six bases of rabbit hepatic lipase coding sequence."  
 Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 [3]  
 SEQUENCE FROM N.A.  
 Warren R.J., Ebert D.L., Mitchell A., Barter P.J.;  
 Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 EMBL; AF041202; AAB96786.1; --  
 HSP; P06857; 1RP1.  
 InterPro; IPR00734; Lipase.  
 InterPro; IPR001024; Lipoxigenase\_LH2.  
 InterPro; IPR000379; Ser\_estra\_site.  
 Pfam; PF00151; lipase; 1.  
 Pfam; PF01477; PLAT; 1.  
 PRINTS; PR00821; TAGLIPASE.  
 SMART; SM00308; LH2; 1.  
 PROSITE; PS00120; LIPASE\_SER; 1.  
 SEQUENCE 499 AA; 55813 MW; 4DA7D6EE815A0AS CRC64;

ary Match 19.8%; Score 501.5; DB 6; Length 499;  
 at Local Similarity 31.1%; Pred. No. 1.1e-34;  
 :ches 140; Conservative 71; Mismatches 192; Indels 47; Gaps 16;  
 54 TRFLYFHNPAYQETISAVNSSTIQASVFGTDKTRINIAWKTDGK-----WORDMCN 108  
 49 TRFLYFK-DRANKGCQRLHADTLQECFNSPLVMTVHGVSDGLLESWQMAAL 107  
 109 VLQLEDINCILNDWIN-GSEYIHAVNRLVVGAEVYFIDVLMKKKEYSKVLHGH 167  
 108 KSQPARPVNGLVDWISLAHSHYAVAVNARLVQGEVALLQWLESAPPFSNVHLIGY 167  
 168 SLGHLAGAGSRIPG---LGRITGLDPAGPFPHNTPEKVRIDPSDANFVDVHNAAARI 224  
 168 SLGHVAGFAGSYISGKHKIGRTGTDAAGPLPEGTASDRILSPDDATFVDAIHT-FTR 226

225 LPFLGYGTIDACCHLDYFVNGGKMPGCE--DLITPLKFNFNAYKKEMASFFPCNHARS 282  
 227 HMLSGVIGIQPVGHVDFYFNGSGFQPGCHFLYKHAHQGLNA-----LSQTIKAHARS 282  
 283 YQFYABSIINPD-AFIATPCRSYTSFKAGNCFPSKGCPTMGHFAHFRHFKNMK-TNGS 340  
 283 VHLFIDSLHFGMSQSTAYQCSMDMSFSQCLCLGCTKGRCNTLG-----YHIROBPLSXGK 337  
 341 HVLNLTGSLSPFAWRHKLKSVKLSG--SEVTQGVFLRVGGAIGKGTGFAIVSGK-LPFG 397  
 338 RFLVYTAQASPRVHYQKFIQFINGIEKRPVPTTMSLLGTKEIKRIPITLGEIGTSN 397  
 398 MYTKLIDADVNGNITSVQFIWK--HLFEDSQNKL-----GAEMVINT-----S 441  
 398 KTVSFLITLNLIDIGELIMVIFKFWNSAVWVWNTVQITLWGIPEHSGILKLTIRVKA 457  
 442 GKYGYKSTFCSDI---MGNILQNLKPC 467  
 458 GETQQRMTFCSENMDDQLHPTQKQVRC 487

PRELIMINARY; PRT; 540 AA.

01-MAY-2000 (TrEMBLrel. 13, Created)  
 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 CG5966 protein (LD47264P).  
 CG5966.  
 O Drosophila melanogaster (Fruit fly).  
 OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Sphingidae; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 [1]  
 SEQUENCE FROM N.A.  
 STRAIN=BERKELEY;  
 MEDLINE=20196006; PubMed=10731132;  
 Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 Brannon R.C., Rogers Y.-H.C., Blazey R.G., Champagne B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 Abril J.F., Achayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
 Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 Reinart K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

tein - protein search, using sw model

n: January 6, 2004, 11:10:48 ; Search time 18 Seconds  
(without alignments)  
1220.082 Million cell updates/sec

st score: 2527

ace: 1 MGIIWVAFLEFGTSGKEV.....STFCSQDINGPNILQNKPC 467

3g table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ted: 127863 seqs, 47026705 residues

number of hits satisfying chosen parameters: 127863

am DB seq length: 0

am DB seq length: 2000000000

processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

ase : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

	Score	Match	Length	ID	Description
1	1191	47.1	470	1 LIP2 MYOCO	Q6424 myocastor c
2	1172	46.4	473	1 LIP1 RAT	P54316 rattus norv
3	1166	46.1	467	1 LIP1 CANFA	P06857 canis famil
4	1162	46.0	467	1 LIP1 HUMAN	P54315 homo sapien
5	1152	45.6	469	1 LIP2 HUMAN	P54317 homo sapien
6	1147.5	45.4	457	1 LIP2 MYOCO	Q6425 myocastor c
7	1142.5	45.2	468	1 LIP2 RAT	P54318 rattus norv
8	1115	44.1	465	1 LIP2 HUMAN	P16233 homo sapien
9	1109	43.9	461	1 LIP2 HORSE	P29183 equus cabal
10	1108	43.8	465	1 LIP2 CAPO	P50903 cavia porce
11	1105	43.7	465	1 LIP2 RABIT	Q02157 oryctolagus
12	1094.5	43.3	468	1 LIP2 MOUSE	P17892 mus musculu
13	1073	42.5	465	1 LIP2 RAT	P27657 rattus norv
14	1069	42.3	450	1 LIP2 PIG	P00591 sus scrofa
15	534	21.1	499	1 LIP2 HUMAN	P11150 homo sapien
16	524	20.7	500	1 LIP2 MOUSE	Q9WY95 mus musculu
17	515.5	20.4	500	1 LIP2 HUMAN	Q9YX59 homo sapien
18	499	19.7	510	1 LIP2 HUMAN	P27656 mus musculu
19	484	19.2	494	1 LIP2 RAT	P07867 rattus norv
20	471	18.6	465	1 LIP2 CAPO	P11153 cavia porce
21	465.5	18.4	475	1 LIP2 MUSVI	Q46647 musccla vis
22	464	18.4	478	1 LIP2 FELCA	P55031 felis silve
23	459	18.2	478	1 LIP2 PIG	P49923 sus scrofa
24	458.5	18.1	490	1 LIP2 CHICK	P11602 gallus gall
25	453	17.9	475	1 LIP2 HUMAN	P06858 homo sapien
26	452	17.9	478	1 LIP2 SHEEP	Q29524 ovis aries
27	450	17.8	475	1 LIP2 PAPAN	Q49060 papio anubi
28	445.5	17.6	474	1 LIP2 RAT	Q06000 rattus norv
29	444.5	17.6	465	1 LIP2 BOVIN	P11151 bos taurus
30	441.5	17.5	474	1 LIP2 MOUSE	P11152 mus musculu
31	291	11.5	301	1 PAL1 MOUSE	Q9U6W0 polistes an
32	287	11.4	303	1 PAL2 DOLMA	P53357 dolichovesp
33	284.5	11.3	300	1 PAL1 VESMC	P51528 vespula mac

p49369 vespula vul  
Q06478 dolichovesp  
P02843 drosophila  
P06607 drosophila  
P02844 drosophila  
P27587 ceratitis c  
P27878 ceratitis c  
Q8Vbx1 rattus norv  
P00688 mus musculu  
Q9ce01 lactococcus  
Q28295 canis famil  
P39533 saccharomyc

34 281 11.1 336 1 PAL1 VESVU  
35 277 11.0 317 1 PAL1 DOLMA  
36 216 8.5 439 1 VIT1\_DROME  
37 192 7.6 420 1 VIT3\_DROME  
38 184.5 7.3 442 1 VIT2\_DROME  
39 183 7.2 422 1 VIT2\_DROME  
40 158.5 6.3 437 1 VIT1\_CERCA  
41 141.5 5.6 80 1 LIPE RAT  
42 97 3.8 508 1 AMYP\_MOUSE  
43 94 3.7 763 1 PEPR\_LACLA  
44 94 3.7 2813 1 VWF\_CANFA  
45 93.5 3.7 789 1 ACOX\_YEAST

#### ALIGNMENTS

RESULT 1  
LIP2\_MYOCO  
ID LIP2\_MYOCO STANDARD; PRT; 470 AA.  
AC Q6424;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Pancreatic lipase related protein 2 precursor (EC 3.1.1.3) (PL-PP2).  
OS Myocastor copys (Coyu) (Nutria).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Myocastoridae;  
OC Myocastor.  
OX NCBI\_TaxID=10157;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-23.  
RC TISSUE=Pancreas;  
RX MEDLINE=95154288; PubMed=7851384;  
RA Thirstrup K., Carriere F., Hjorth S.A., Rasmussen P.B., Nielsen P.F.,  
Ladefoged C., Thim L., Boel E.;  
RT "Cloning and expression in insect cells of two pancreatic lipases and  
a procolipase from Myocastor copys.";  
RL Eur. J. Biochem. 227:186-193 (1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94176463; PubMed=8130186;  
RA Thirstrup K., Verger R., Carriere F.;  
RT "Evidence for a pancreatic lipase subfamily with new kinetic  
properties.";  
RL Biochemistry 33:2748-2756 (1994).  
CC -! CATALYTIC ACTIVITY: triacylglycerol + H(2)O = diacylglycerol + a  
fatty acid anion.  
CC -! TISSUE SPECIFICITY: Pancreas.  
CC -! SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.  
CC -! SIMILARITY: Contains 1 PLAT domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X83000; CAA58121.1; -.  
DR PIR; A54232; A54232.  
DR HSSP; P54318; 1BU8.  
DR InterPro; IPR000734; Lipase.  
DR InterPro; IPR001024; Lipoxigenase LH2.  
DR InterPro; IPR000379; Ser\_estra\_site.  
DR Pfam; PF00151; lipase; 1.  
DR Pfam; PF01477; PLAT; 1.  
DR PRINTS; PR00821; TAGLIPASE.  
DR SMART; SM00308; LH2; 1.  
DR PROSITE; PS00120; LIPASE\_SER; FALSE\_NEG.  
DR PROSITE; PS50095; PLAT; 1.  
KW Hydrolase; Lipid degradation; Pancreas; Glycoprotein; Signal.

1 SIGNAL 18  
CHAIN 19 470 PANCREATIC LIPASE RELATED PROTEIN 2.  
DOMAIN 358 470 PLAT.  
ACT\_SITE 172 196 CHARGE RELAY SYSTEM (BY SIMILARITY).  
ACT\_SITE 196 196 CHARGE RELAY SYSTEM (BY SIMILARITY).  
ACT\_SITE 283 283 CHARGE RELAY SYSTEM (BY SIMILARITY).  
DISULFID 22 28 BY SIMILARITY.  
DISULFID 110 121 BY SIMILARITY.  
DISULFID 257 281 BY SIMILARITY.  
DISULFID 305 316 BY SIMILARITY.  
DISULFID 319 324 BY SIMILARITY.  
DISULFID 454 470 BY SIMILARITY.  
SEQUENCE 470 AA; 52060 MW; 5F06DBC8F64869 CRC64;

ary Match 47.1%; Score 1191; DB 1; Length 470;  
3c Local Similarity 49.8%; Pred. No. 6.6e-91;  
::ches. 236; Conservative 63; Mismatches 163; Indels 12; Gaps 7;

1 MLGIWIVAFLLFFGTSGKVCYERLCGPKDGLPWTRTFSTELVGLPWSPEKINIRFLIYT 60  
2 MLFWTTGGLLLLTATAGNEVCYSHLGCSDKPKAGTLQRPVKSLPASPEINIRFLIYT 61  
61 IHNPNVQEIISAVNSSTIOASYFGTDKTIIRINIAGWKTGK--WORDMNCVLLQJEDINC 118  
62 NENPNYQLITADPATIRASFNLRHXTRFVIHGFIDNGEKDMLTDICKRMFQVEKVC 121  
119 INLDWINGSRE-YIHAVNNLVVGAEVAYFDVLWKPFYSPSKVHLCHSLGAHLAGEA 177  
122 ICVDWQGSGLIYSQAVQNIQVVGAEVAYLVQVLSLDOLGYPGNVHMIGHSLGHTAAEA 181  
178 GSRIPGL-GRITGLDPAGFFHNTPEKVRLPDSANPDVDVIHTNARILFELGVGTIDAC 236  
182 GRRLKGLVGRITGLDPAEPCPDTEEVRLDPDSAMFVDVIHTDIAPILPSFGFGMSQKV 241  
237 GHLDYFNGKHNPGCE-DLITPLKFNFNAYKEMASFPDCHNARSYQFAESILNPDA 295  
242 GHMDFPFGKHPGCEKNIISTIV--DVNGFLEGITSLAACNHRYSQYSSSILNPDG 299  
296 FIAPCPSRYTSFKAGNCPFCSEKGGOPTMGHPADRFHFGMKNTG--SHYFLNTGSLSPA 353  
300 FLGYPCASYEEFKDCCFPACPGCKVGHVADQF---QKGANGVEKYFYLTNGDSNFP 356  
354 RWRKLSVKLUSGEVITQGTVFLRVGGAIGKTGEFAIVSGKLEPGMTYTKLIDADVNGNI 413  
357 RWRKYSVTLSGEKELSGDGIKIALFGRNGSKQVEIFKGLSKPDARYTHDIDVLNVGEI 416  
414 TSVQFIKKHLFEDSQNLGAEMWINTSGKYGKSTFCSDINGFNILQNLKPC 467  
417 QKVAFLLHNNGINLLOPKLGASQITVQSGBYGTYKYNFCSSNTVQEDVLQSLSPC 470

JT 2  
RAT  
LIP1 RAT  
P54316; STANDARD; PRT; 473 AA.  
01-OCT-1996 (Rel. 34, Created)  
01-OCT-1996 (Rel. 34, Last sequence update)  
15-SEP-2003 (Rel. 42, Last annotation update)  
Pancreatic lipase related protein 1 precursor (EC 3.1.1.33).  
PNLIPR1 OR FLRPI.  
Rattus norvegicus (Rat).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
NCBI\_TaxID=10116;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=Wistar; TISSUE=Pancreas;  
MEDLINE=92111786; PubMed=1730292;  
Wicker-Planquart C., Puigserver A.;  
"Primary structure of rat pancreatic lipase mRNA."  
FEBS Lett. 296:61-66(1992).  
-!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a  
fatty acid anion.

```
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
CC -!- SIMILARITY: Contains 1 PLAT domain.
CC -!- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE THE PANCREATIC
CC      LIPASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X61925; CAA43927.1; -.
CC PIR; S20612; S20612.
CC HSSP; P06857; IRP1.
CC InterPro; IPR001074; Lipase.
CC InterPro; IPR001024; Lipoxigenase LH2.
CC InterPro; IPR00379; Ser_estr_site.
CC Pfam; PF00151; lipase; 1.
CC PRINTS; PF01477; PLAT; 1.
CC SMART; SM00308; TAGH2; 1.
CC PROSITE; PS00120; LIPASE_SER; 1.
CC PROSITE; PSS0095; PLAT; 1.
CC Hydrolase; lipid degradation; Pancreas; Signal.
KW SIGNAL
FT CHAIN          1..17 POTENTIAL.
FT COMPACT        1..17
FT DOMAIN         356..470 PLAT.
FT ACT_SITE       171..171
FT ACT_SITE       194..194
FT ACT_SITE       281..281
FT ACT_SITE       281..281
FT DISULFID       21..27
FT DISULFID       109..120
FT DISULFID       255..279
FT DISULFID       303..314
FT DISULFID       317..322
FT DISULFID       451..467
FT CARBOHYD       157..157 N-LINKED (GLCNAC...) (POTENTIAL).
FT SEQUENCE       473 AA; 52378 MW; C03DIE021F30723 CRC64;
Query Match              46.4%; Score 1172; DB 1; Length 473;
Best Local Similarity   49.3%; Pred. No. 2.5e-89;
Matches 232, Conservative 73; Mismatches 158; Indels 8; Gaps 6;
QY    1  MLGIWIVAFELFGTSRGKVCYERLGCGFDGLPWTRFSTFELVGLPMSPEKINTRLIYT 60
DB    1  MLTLWTVSFLGAAQGKCEVDNIGCFSDAEPWAGTAIRPLKLLPMSPEKINTRLIYT 60
QY    61  IHNPNAVQISAVNSGTQASVFQDKTKRINIAGW--KTDGKWDQMDCNVLQLLEDINC 118
DB    61  NENPATAQTQLSLDLPTICASNFOYARKTRIIHFDTKGEEVNVDYDMCKMFOVEVNC 120
QY    119  INLDWNGSR-EYIHAVNNLRVGAEVAFTDVLMKKFEYSPSKVHLHGSHLAHLAGEA 177
DB    121  ICVDWKKSQTTVTOAANNVRVYGQAQMIDIIVKNYSYSPSKVHLHGSHLAHVAGEA 180
QY    178  GSRIPLGLGRITGLDPAGPFPHNTPKVELDFDSANFVDVIHTHAARTLPFLGVGTIDAOG 237
DB    181  GSTPPLGLRITGLDPVEANFGTPPEVRLDSPDAFVDVIHTDAAPLIIFLPGFTNQMSG 240
QY    238  HLDFFPNGGKHMPGC-EDLTITPLLKFNFAYKEMASFPDCNHARSYQFYAESILNPDFAF 296
DB    241  HLDFFPNGGQSPGCKNALSQIV--DIDGIWSGTDRFVACNELRSYKYLYESILNPDGF 298
QY    297  IAYPCSYTSFRAGNCFCSKEGCPCTMGCHPADRFHEFNKMKTNGSHVPLNTGSLSPEARWE 356
DB    299  AAIPCSAYKDFSNKCFCPCDQCQGMGHYADKFAKGS-GDEPKCFPLNTGEANKFARWR 357
QY    357  HKLSVKLSGSEVTQGTVFILRVGCAIGKTGEPAIVSGKLEPGMTVTKLIDADVNVGNITSV 416
DB    358  VRYSLIJSGRWIT-GOVKVAFSGKGNTRPVDIIFGLIKKGATHSSSEFAKLVDGTVIKV 416
```





STRAND 448 451  
 STRAND 455 456  
 TURN 458 459  
 STRAND 462 467  
 SEQUENCE 467 AA; 51482 MW; 531868454048535 CRC64;  
 try Match 46.18; Score 1166; DB 1; Length 467;  
 Local Similarity 48.0%; Pred. No. 7.7e-89;  
 Cons 226; Conservative 79; Mismatches 158; Indels 8; Gaps 6;  
 1 MLGIWVAFPLFFGTSGKVCYERLGLCPDGLPWTFTSTELVGLPWSPEKINTRELLYT 60  
 1 MVSITIALFLLGAARAKAVCYQICGSDAEPAWTAIRPLKVLFPWSPERIGTRFLYT 60  
 61 IHNPNAYOEISAVNSSTIOASYFGTDKITRINIAGW--KTDGKQWDMCNVLQLEIDNC 118  
 61 NKNPNFQTLPSDPTIEASNFQTKTRFIHGFIDKGEENWLLDMCKMFKVEEVNC 120  
 119 INLDWINGSR-EYTHAVNNLRVGAEVAFIDVLMKKFYSKVLHLSGHAHLAGEA 177  
 121 ICVDWKKGSQATYQAANNRVGAQVQWMLSNYSYFQVQLHLSGHAHLAGEA 180  
 178 GSRIPLGLRITGLDPAAGPFPHNTPEKVRLOPDSANFVDVHTNAARILFELGVGTIDACG 237  
 181 GSRTPEGLRITGLDPAEASFGTPEEVRLDPTDADFDVHTDAAPLIPFLGFGTSQQWG 240  
 238 HLDYPNGKEMPGC-EDLITPLKFNFNAYKEMASFPDCHNARSYQFYAESILNDPAF 296  
 241 HLDFFPNGGEMPGCKKNAISQIV--DLDGWEGTRDFVACNHLRSYKYSESILNDPGF 298  
 297 IAYPCRSYTSFKAGNCFKSCGCTPMGCFADRFHFNKMTNGSHYFNTGSLSPFARWR 356  
 299 ASYPCASYRAFESKCFPCPDQCPQWGHVADKFAVKT-SDTQKYFLNTGSSNFARWR 357  
 357 HKLSVKLSGSEVTOGTFTLVGGAIGTGEFAIVSGKLEPQMTYTKLIDADVNGNITSV 416  
 358 YGVSIITLSGRTAT-GQIKVALFGNKGNTHQYSIFRGLKPGSTHSYEFDAKLDVGTIEKV 416  
 417 QFTWKKHLFEDSQNKLGAEMVINTSGKYGKSTFCSDIMGNPNILQNLKPC 467  
 417 KFLWNNVNPFPFKYGAAKITVQGEETVHSFSESTVREDVLLITLTPC 467

AT 4  
 HUMAN  
 LIPI HUMAN STANDARD; PRT; 467 AA.  
 P54315;  
 01-OCT-1996 (Rel. 34, Created)  
 01-OCT-1996 (Rel. 34, Last sequence update)  
 15-SEP-2003 (Rel. 42, Last annotation update)  
 Pancreatic lipase related protein 1 precursor (EC 3.1.1.3).  
 PNLI1P1 OR PLRP1.  
 Homo sapiens (Human).  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TaxID=9606;  
 [1]  
 SEQUENCE FROM N.A.  
 TISSUE=Pancreas;  
 MEDLINE=92355622; PubMed=1379598;  
 Giller T., Buchwald P., Blum-Kaelin D., Hunziker W., hPLRP1 and  
 hPLRP2, differences in colipase dependence and in lipase activity,"  
 J. Biol. Chem. 267:16509-16516(1992).  
 -!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a  
 fatty acid anion.  
 -!- SUBCELLULAR LOCATION: Secreted (Potential).  
 -!- TISSUE SPECIFICITY: Pancreas.  
 -!- SIMILARITY: BELONGS TO THE ABHDROLASE SUPERFAMILY. LIPASE FAMILY.  
 -!- SIMILARITY: Contains 1 PLAT domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation.

CC the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; M93283; AAS9532.1; -  
 CC PIR; A43357; A43357.  
 CC DR HGPS; P06857; LRPI.  
 CC DR Genew; HGNC:9156; PNLI1P1.  
 CC DR MIM; 604422; Cextracellular; TAS.  
 CC DR GO; GO:0005576; Cextracellular; TAS.  
 CC DR GO; GO:0004806; Fittacylglycerol lipase activity; TAS.  
 CC DR InterPro; IPR000734; Lipase.  
 CC DR InterPro; IPR001024; Lipoxigenase LH2.  
 CC DR InterPro; IPR000379; Ser estersite.  
 CC DR Pfam; PF00151; lipase; 1-  
 CC DR Pfam; PF01477; PLAT; 1-  
 CC DR PRINTS; PRO0821; TAGLIPASE.  
 CC DR SMART; SM00308; LH2; 1.  
 CC DR PROSITE; PS00120; LIPASE SER; 1.  
 CC DR PROSITE; PS00095; PLAT; 1.  
 CC DR Hydrolase; Lipid degradation; Pancreas; Signal; Polymorphism.  
 CC FT CHAIN 1 17  
 CC FT SIGNAL 1 17  
 CC FT CHAIN 18 467  
 CC FT DOMAIN 356 467  
 CC FT ACT\_SITE 171 171  
 CC FT ACT\_SITE 194 194  
 CC FT ACT\_SITE 281 281  
 CC FT DISULFID 21 27  
 CC FT DISULFID 109 120  
 CC FT DISULFID 255 279  
 CC FT DISULFID 303 314  
 CC FT DISULFID 317 322  
 CC FT DISULFID 451 467  
 CC FT VARIANT 461 461  
 CC FT L -> P (IN DBSNP:1049125).  
 CC FT /FTID=VAR 014915.  
 CC SQ SEQUENCE 467 AA; 51847 MW; 2781EA1E22E39E78 CRC64;  
 Query Match 46.0%; Score 1162; DB 1; Length 467;  
 Best Local Similarity 48.6%; Pred. No. 1.6e-88;  
 Matches 229; Conservative 75; Mismatches 159; Indels 8; Gaps 6;  
 QY 1 MLGIWVAFPLFFGTSGKVCYERLGLCPDGLPWTFTSTELVGLPWSPEKINTRELLYT 60  
 DB 1 MLFIWITLFLGAAKGEVCYEDLGCFSDETPGWTARPLKILPWSPEKIGTRFLYT 60  
 QY 61 IHNPNAYOEISAVNSSTIOASYFGTDKITRINIAGW--KTDGKQWDMCNVLQLEIDNC 118  
 DB 61 NKNPNFQTLPSDPTIEASNFQTKTRFIHGFIDKGEENWLLDMCKMFKVEEVNC 120  
 QY 119 INLDWINGSR-EYTHAVNNLRVGAEVAFIDVLMKKFYSKVLHLSGHAHLAGEA 177  
 DB 121 ICVDWKKGSQATYQAANNRVGAQVQWMLSNYSYFQVQLHLSGHAHLAGEA 180  
 QY 178 GSRIPLGLRITGLDPAAGPFPHNTPEKVRLOPDSANFVDVHTNAARILFELGVGTIDACG 237  
 DB 181 GSRTPEGLRITGLDPAEASFGTPEEVRLDPTDADFDVHTDAAPLIPFLGFGTSQQWG 240  
 QY 238 HLDYPNGKEMPGC-EDLITPLKFNFNAYKEMASFPDCHNARSYQFYAESILNDPAF 296  
 DB 241 HLDFFPNGGEMPGCKKNAISQIV--DLDGWEGTRDFVACNHLRSYKYSESILNDPGF 298  
 QY 297 IAYPCRSYTSFKAGNCFKSCGCTPMGCFADRFHFNKMTNGSHYFNTGSLSPFARWR 356  
 DB 299 ASYPCASYRAFESKCFPCPDQCPQWGHVADKFAVKT-SDTQKYFLNTGSSNFARWR 357  
 QY 357 HKLSVKLSGSEVTOGTFTLVGGAIGTGEFAIVSGKLEPQMTYTKLIDADVNGNITSV 416  
 DB 358 YGVSIITLSGRTAT-GQIKVALFGNKGNTHQYSIFRGLKPGSTHSYEFDAKLDVGTIEKV 416  
 QY 417 QFTWKKHLFEDSQNKLGAEMVINTSGKYGKSTFCSDIMGNPNILQNLKPC 467

417 KFLMNNVINTPLKVGATKITVQGEKTVYNFCSSEDTVRDITLLTTPC 467

HUMAN  
LIP2 HUMAN STANDARD; PRT; 469 AA.  
PS4317;  
01-OCT-1996 (Rel. 34, Created)  
01-OCT-1996 (Rel. 34, Last sequence update)  
15-SEP-2003 (Rel. 42, Last annotation update)  
Pancreatic lipase related protein 2 precursor (EC 3.1.1.3).  
PNLIPR2 OR PLRP2.  
Homo sapiens (Human).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
[1]  
NCBI\_TaxID=9606;  
SEQUENCE FROM N.A.  
TISSUE=Pancreas;  
MEDLINE=92355622; PubMed=1379599;  
Galler T., Buchwald P., Blum-Kaelin D., Hunziker W.;  
"Two novel human pancreatic lipase related proteins, hPLRP1 and  
hPLRP2. Differences in colipase dependence and in lipase activity.";  
J. Biol. Chem. 267:16509-16516(1992).  
[2]  
SEQUENCE FROM N.A.  
TISSUE=Pancreas;  
MEDLINE=22388257; PubMed=12477932;  
Strausberg R.L., Feigold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
Brownstein M.J., Uedin T.B., Teshiyouki S., Carninci P., Prange C.,  
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
Besak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalón D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
"Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences.";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
-!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a  
fatty acid anion.  
-!- TISSUE SPECIFICITY: Pancreas.  
-!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.  
-!- SIMILARITY: Contains 1 PLAT domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M93284; AAAS5933.1;  
EMBL; BC005989; AAH05989.1;  
PIR; B43357; B43357.  
HSP; P54318; 1BUB.  
Gene; HGNC:9157; PNLIPR2.  
MIM; 604423;  
GO; GO:0005576; C:extracellular; TAS.  
GO; GO:0004806; F:triacylglycerol lipase activity; TAS.  
GO; GO:0006641; P:triacylglycerol metabolism; TAS.  
InterPro; IPR000734; Lipase.  
InterPro; IPR001024; Lipoxigenase\_LH2.

DR InterPro; IPR000379; Ser\_estrs\_site.  
DR Pfam; PF00151; lipase; 1.  
DR Pfam; PF01477; PLAT; 1.  
DR PRINTS; PRO0821; TAGLIPASE.  
DR SMART; SM00308; LH2; 1.  
DR PROSITE; PS00120; LIPASE\_SER; 1.  
DR PROSITE; PS00095; PLAT; 1.  
KW Hydrolase; Lipid degradation; Pancreas; Glycoprotein; Signal.  
FT SIGNAL 1 17  
FT CHAIN 18 469  
FT DOMAIN 357 469  
FT ACT\_SITE 171 171  
FT ACT\_SITE 195 195  
FT ACT\_SITE 282 282  
FT DISULFID 21 27  
FT DISULFID 109 120  
FT DISULFID 256 280  
FT DISULFID 304 315  
FT DISULFID 318 323  
FT DISULFID 453 469  
FT CARBOHYD 353 353  
FT CARBOHYD 428 428  
SQ SEQUENCE 469 AA; 51946 MW; DOEL130295A94A725 CRC64;  
Query Match 45.6%; Score 1152; DB 1; Length 469;  
Best Local Similarity 47.8%; Pred. No. 1.1e-87;  
Matches 225; Conservative 68; Mismatches 172; Indels 6; Gaps 5;  
QY 1 MLGIVIAFLPFGTSGRGVCVERLGCPCDGLPWTRFSTELVGLPWSPEKINTRFLLYT 60  
DB 1 MLPPWTGLLLATVGRGEVYGGQCSDEKPKAGTLQRFVKLLPWSPEIDTRFLLYT 60  
QY 61 IHNPAYQEISAVNSSTIOASYFGTDKITRINIAGW--KTDGKWRDMCNVLLQLEDINC 118  
DB 61 NENPNFOLITGTEPTIEAGNFOLDKTRFIIHGFLDKAEDSWESDMCKMKEVEKVCNC 120  
QY 119 INLDWINGSRE-YIHANNVNLVGAEVAYFDVLVKFEYSPSKVHLIGHSLGAHLAGEA 177  
DB 121 ICVDRHGRSRAMYTQAVQINRVGAETAFLQALSTQLGVSLESDVHVHIGSLGAHTAAEA 180  
QY 178 GSRIPG-LGRITGLDPAGFPFHNTKPEVRLDPSDANFVDVHTNAARILFELGVGTIDAC 236  
DB 181 GRRLGRVGRITGLDPAGFCQDSEEVRLDPSDAVFDVHTDSSPIVPSLGFMSQKV 240  
QY 237 GHLDYPNGGKMGCCEDLITPLKFNFNAYKEMASFFDCNHAARSQFYAESILNPDFA 296  
DB 241 GHLDYFPNGGKMGCKKNVLSITIT-DIDGIGWEGIGGVSCNHLRSEFYSSVLPDGF 299  
QY 297 IAYPCRSYTSFKAGNCPFCSEKGCPTMGHPADRFHFKNMKNGSHVFLNTGSLSPFARWR 356  
DB 300 LGYPCASVDETSQSKCPCPAEGCPKMGHYADQFKGKTSAVEQT-FFLNTGESGNFTSWR 358  
QY 357 HKLSVKLSGSEVTOGTVFLRVGGAIGKTGTFPAIVSGKLPEGMYTYTKLIDADVNGNITSV 416  
DB 359 YKVSVTLTGKWKVGYRIALYGSNENSKQVEIFKGLKPDASHTCAIDVDVFNVGKIQV 418  
QY 417 QPIWKHLFPDSQNKLGAEWINTSGYGVKSTFCSDODINGPILQNLKPC 467  
DB 419 KFLMNNKGINLSKPLGASQITVOSGEDGTEYFNFCSSDTEVENVLOSLYPC 469

RESULT 6  
LIPP MYOOC  
ID LIPP MYOOC STANDARD; PRT; 457 AA.  
AC Q64425;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Triacylglycerol lipase, pancreatic precursor (EC 3.1.1.3) (Pancreatic  
lipase) (PL) (Fragment).  
GN PNLIP  
OS Myocastor corpus (Coyup) (Nurria).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Rodentia; Hystricognathi; Myocastoridae;  
Myocastor.

NCBI\_TaxID=10157;

[1]

SEQUENCE FROM N.A.

TISSUE=Pancreas;

MEDLINE=9515488; PubMed=7851384;

Thisstrup K., Carriere F., Hjorth S.A., Rasmussen P.B., Nielsen P.F.,

Ladefoged C., Thim L., Boel E.;

"Cloning and expression in insect cells of two pancreatic lipases and

a procolipase from Myocastor coypus.";

Eur. J. Biochem. 227:186-193 (1995).

-!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a

fatty acid anion.

-!- SUBCELLULAR LOCATION: Secreted.

-!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.

-!- SIMILARITY: Contains 1 PLAT domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

EMBL: X82959; CAA58120.1; -.

PIR: I48206; I48206.

HSSP: P29183; LHPL.

InterPro: IPR000734; Lipase.

InterPro: IPR001024; Lipoxigenase LH2.

InterPro: IPR000379; Ser\_estrs\_site.

Pfam: PF00151; lipase; 1.

Pfam: PF01477; PLAT; 1.

PRINTS: PR00821; TAGLIPASE.

SMART: SM00308; LH2; 1.

PROSITE: PS00120; LIPASE SR; 1.

PROSITE: PS00095; PLAT; 1.

Hydrolase; Lipid degradation; Pancreas; Glycoprotein; Signal.

NON\_TER 1 1

SIGNAL <1 7

CHAIN 347 457

DOMAIN 347 457

ACT\_SITE 161 185

ACT\_SITE 185 185

ACT\_SITE 272 272

DISULFID 12 18

DISULFID 99 110

DISULFID 246 270

DISULFID 294 305

DISULFID 308 313

DISULFID 441 457

CARBOHYD 343 343

SEQUENCE 457 AA; 50209 MW; 33D572B770459DAD CRC64;

ery Match

st Local similarity 45.4%; Score 1147.5; DB 1; Length 457;

atches 232; Conservative 69; Mismatches 149; Indels 15; Gaps 10;

10 LFFGTSRGKVEYRLGCGKDLPTWTFSTELVGLPWSPEKINTRELLYTHNPAYQE 69

1 LLLGAVASEVYDLRGCFSDSPAGIVRPLKVLWPSPTINTRELLYTHNPAYQI 60

70 ISAVNSSTIQASYFGTKITRINAGW--KTDGKWORDMCMVLLQLDINCINLDWINGS 127

61 VTA-DSSTIRSNFRTRKTRFIHGYIDKGEENLANMCEALLQVESVNCICVDNMGKS 119

128 RE-YIHAVNVLWCAVEAYFDVLKFFYSPKVLHIGSLGAHLGAGSRIG-LG 185

120 RALYQATONIRWGAWEAYFDALQSGOLGSPSNVHIGSLGSHVAGSAGRTNGTG 179

186 RITGLDPAEPFPHNTPEKVRDPSDANPDVVIHNAARILFELGVGTIDACGHLDPYNG 245

Db 180 RITGLDPAEPFPHNTPEKVRDPSDANPDVVIHNAARILFELGVGTIDACGHLDPYNG 239

Qy 246 GKHPGCE-DLITPLKFNENAYKEMASFFDCNCHARSYQFYAESILNPDATAYPCRY 304

Db 240 GVEMPGCKNIISQIV--DINGIWEGRTRDFAACNHLRSYKYIYDSILNPTGFAGFSCSY 297

Qy 305 TSFKAGNCFKSCSGCPTMGHFAFRHFKNMKTNG--SHYFLNTGSLSPFARWRHKLVS 362

Db 298 NTFSSNCFPCASGCGPQMGHYADRF--SGKTNELFQQFYLTNGDASNFRWRYQIATV 354

Qy 363 LSGSEVTCGTVFLVGAIGKTEGFAISVGLKLPQMGTYKLIDADVNGNITSVQFIWK 422

Db 355 LSGHKVT-GHVLVSLYSGSGTSKYEIKYKSLQFGTSYVNQIDSDVDVGDIEKVKFIW 413

Qy 423 HLFEDSQNKLGAEWVINTSGYKVKSTFCQSDMGPNILQNLKPC 467

Db 414 NIINPTLFKVGASS-IQVTRDGRVNFCSQDVRREDILLTLTPC 457

# RESULT 7

LIP2\_RAT ID LIP2\_RAT STANDARD; PRT; 468 AA.

AC P54318; 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Pancreatic lipase related protein 2 precursor (EC 3.1.1.3) (Secretory

DE Glycoprotein GP-3)

GN PNLIPR2 OR PLRP2.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 17-46.

RC STRAIN=Sprague-Dawley; TISSUE=Pancreas;

RX MEDLINE=93252914; PubMed=8486693;

RA Wishart M.J., Andrews P.C., Nichols R., Blevins G.T. Jr.,

RA Logsdon C.D., Williams J.A.;

RT "Identification and cloning of GP-3 from rat pancreatic acinar

RT zymogen granules as a glycosylated membrane-associated lipase.";

RL J. Biol. Chem. 268:10303-10311(1993).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=94262798; PubMed=8203536;

RA Payne R.M., Sims H.P., Jennens M.L., Lowe M.E.;

RT "Rat pancreatic lipase and two related proteins: enzymatic properties

RT and mRNA expression during development.";

RL Am. J. Physiol. 266:G914-G921(1994).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 17-468.

RC TISSUE=Pancreas;

RX MEDLINE=99041983; PubMed=9822688;

RA Roussel A., Yang Y., Ferrato F., Verger R., Cambillau C., Lowe M.;

RT "Structure and activity of rat pancreatic lipase-related protein 2.";

RL J. Biol. Chem. 273:32121-32128(1998).

CC -!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a

CC fatty acid anion.

CC -!- SUBCELLULAR LOCATION: SECRETED OR MEMBRANE-ASSOCIATED

CC -!- TISSUE SPECIFICITY: PANCREAS. PANCREATIC SECRETORY (ZIMOGEN)

CC GRANULE.

CC -!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.

CC -!- SIMILARITY: Contains 1 PLAT domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

CC -----

CC EMBL: L09216; AAA41250.1; ALT\_INIT.

PDB; 1BU8; 23-DEC-98.  
 InterPro; IPR000734; Lipase.  
 InterPro; IPR001024; Lipoxigenase LH2.  
 InterPro; IPR000379; Ser\_estra\_site.  
 Pfam; PF00151; lipase; 1.  
 Pfam; PF01477; PLAT; 1.  
 PRINTS; PR00821; TAGLIPASE.  
 SMART; SMO0308; LH2; 1.  
 PROSITE; PS00120; LIPASE\_SER; 1.  
 PROSITE; PS00095; PLAT; 1.  
 Hydrolase; Lipid degradation; Pancreas; Glycoprotein; Signal;  
 Membrane; 3D-structure.  
 SIGNAL; 1 16  
 CHAIN 17 468  
 DOMAIN 356 468  
 ACT\_SITE 170 170  
 ACT\_SITE 194 194  
 ACT\_SITE 281 281  
 DISULFID 20 26  
 DISULFID 108 119  
 DISULFID 255 279  
 DISULFID 303 314  
 DISULFID 317 322  
 DISULFID 452 468  
 CARBOHYD 352 352  
 STRAND 18 20  
 HELIX 22 24  
 STRAND 26 28  
 TURN 31 32  
 STRAND 36 36  
 TURN 37 38  
 STRAND 39 39  
 HELIX 48 51  
 TURN 52 52  
 STRAND 54 59  
 TURN 60 61  
 STRAND 62 69  
 TURN 74 74  
 HELIX 75 78  
 TURN 79 79  
 TURN 84 85  
 STRAND 87 92  
 TURN 98 99  
 TURN 101 102  
 HELIX 103 112  
 TURN 113 114  
 STRAND 117 123  
 HELIX 125 128  
 HELIX 132 157  
 HELIX 161 163  
 STRAND 164 169  
 TURN 170 170  
 HELIX 171 182  
 TURN 183 185  
 STRAND 189 193  
 STRAND 196 196  
 TURN 198 202  
 HELIX 205 207  
 HELIX 211 213  
 STRAND 217 220  
 HELIX 227 230  
 STRAND 234 234  
 STRAND 242 246  
 TURN 247 248  
 TURN 253 254  
 HELIX 266 270  
 TURN 271 272  
 HELIX 279 283  
 HELIX 295 297  
 STRAND 299 299  
 STRAND 301 302  
 HELIX 306 310  
 TURN 311 312

N-LINKED (GLCNAC. . .) (POTENTIAL).

FT TURN 319 320  
 FT STRAND 324 324  
 FT HELIX 327 331  
 FT TURN 333 336  
 FT STRAND 341 345  
 FT TURN 351 352  
 FT STRAND 356 365  
 FT STRAND 370 380  
 FT TURN 381 382  
 FT STRAND 383 384  
 FT STRAND 388 395  
 FT TURN 397 398  
 FT STRAND 400 407  
 FT STRAND 414 422  
 FT STRAND 433 442  
 FT TURN 443 445  
 FT STRAND 448 452  
 FT STRAND 457 457  
 FT TURN 459 460  
 FT STRAND 463 468  
 SQ SEQUENCE 468 AA; 52534 MW; B41F80B339BA9A6F CRC64;

Query Match 45.2%; Score 1142.5; DB 1; Length 468;  
 Best Local Similarity 48.6%; Pred. No. 6.7e-87;  
 Matches 229; Conservative 63; Mismatches 172; Indels 7; Gaps 6;

QY 1 MGIWIVAFPFGRSGKEVCYERLGCDFKDLGFWTRTFSTELVGLPWSPEKINTRFLLYT 60  
 DB 1 MLLCWIVS-LLLATVGGKEVCYGHLCGFSNDKPKWAGMLQRLKIFPWSPEIDITRFLLYT 59  
 QY 61 IHNPNAYQEISAVNSSTIQASYFGTDXITRINAGW--KYDGKQWDMCNVLLQLEDINC 118  
 DB 60 NENPNYQKISATEPDTIKFSNFDLDRKTRFI VHGFIDKGEDGMLLDMCKMPQVEKVC 119  
 QY 119 INLDWINGSR-EYIHAVNLRVVGAEVAFIDVLMKKFEYSPSKVHLIGHSLGAHLAGEA 177  
 DB 120 ICVDMRGSRTEYTCASVNTFVVGAEAFVQLVSTEMGYSPENVHLIGHSLGAHVGEA 179  
 QY 178 GSRIPG-LGRITGLDPAGPFHNTPKVRLDPSDANFVDVHTNAARILFELGVGTIDAC 236  
 DB 180 GRLLEGHVGRITGLDPAEPFCQGLPBEVRLDPSDANFVDVHTDSAPIIPYLGFGMSQKV 239  
 QY 237 GHLDYFNGGKHPGCCDLITPLLKFNFNAYKEMASFPDCNHSYQVFAESILNPDAF 296  
 DB 240 GHLDFFNGGKHPGCCOKNLSTI-VDINGIWEGTQNFVACNHLRSYKYIASSILNPDF 298  
 QY 297 IAYPCRSYTSFKAGNCFFCSKGCPTMGHFAFRHFKNMKTNGSHYFLNTGSLSPFARWR 356  
 DB 299 LGYPCSSYEFQNDCCPCEGCPKMGHVADQFEGKTATVEQTVY-LNTGDSGNFTWR 357  
 QY 357 HKLSVKLSGSEVTQGTVFLRVGGAIGKTGEFALVSGKLEPGMTYTKLIDADVNVGNITSV 416  
 DB 358 YKSVTLUSGAKLSGYILVALYGNNGSKQYEIFKGLKPEARHVRDIDVDINVGEIQKV 417  
 QY 417 QFIWKHLFEDSQNLGAEMVINTSGKYGVKSTFCSDIMGNILQNLKPC 467  
 DB 418 KFLWNKVINLFRPTILGASQITVQSGVDGKYNFCSSDVRDVLQSLYPC 468

## RESULT 8

LIPP HUMAN  
 ID LIPP HUMAN STANDARD; PRT; 465 AA.  
 AC P16233;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Triacylglycerol lipase, pancreatic precursor (EC 3.1.1.3) (Pancreatic  
 DE lipase) (PL).  
 GN PNLIP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;

[1] SEQUENCE FROM N.A.  
MEDLINE=90062115; PubMed=2479644;  
Lowe M.E., Rosenblum J.L., Straus A.W.;  
"Cloning and characterization of human pancreatic lipase cDNA."; J. Biol. Chem. 264:20042-20048 (1989).

[2] SEQUENCE FROM N.A.  
TISSUE=Pancreas;  
MEDLINE=92355622; PubMed=1379598;  
Giller T., Buchwald P., Blum-Kaelin D., Hunziker W.;  
"Two novel human pancreatic lipase related proteins; hPLRP1 and hPLRP2. Differences in colipase dependence and in lipase activity."; J. Biol. Chem. 267:16509-16516 (1992).

[3] SEQUENCE FROM N.A.  
MEDLINE=94010322; PubMed=8406023;  
Sims H.P., Jennens M.L., Lowe M.E.;  
"The human pancreatic lipase-encoding gene: structure and conservation of an Alu sequence in the lipase gene family."; Gene 131:281-285 (1993).

[4] SEQUENCE FROM N.A.  
TISSUE=Lung;  
MEDLINE=22388257; PubMed=12477937;  
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Staptchuk M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Tohivukki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Heiton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Wyers R.M., Butlerfield J.S.N., Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

[5] X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
MEDLINE=90158821; PubMed=2106079;  
Winkler F.K., D'Arcy A., Hunziker W.;  
"Structure of human pancreatic lipase."; Nature 343:771-774 (1990).

[6] X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).  
MEDLINE=92396238; PubMed=1522902;  
van Tilbeurgh H., Sarda L., Verger R., Cambillau C.;  
"Structure of the pancreatic lipase-procolipase complex."; Nature 359:159-162 (1992).

[7] X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).  
MEDLINE=93241293; PubMed=8479519;  
van Tilbeurgh H., Egloff H.-P., Martinez C., Rugani N., Verger R., Cambillau C.;  
"Interfacial activation of the lipase-procolipase complex by mixed micelles revealed by X-ray crystallography."; Nature 362:814-820 (1993).  
-!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a fatty acid anion.  
-!- SUBCELLULAR LOCATION: Secreted.  
-!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.  
-!- SIMILARITY: Contains 1 PLAT domain.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation.

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC DR EMBL; J05125; AAA36740.1; -  
CC DR EMBL; M93285; AAA60129.1; -  
CC DR EMBL; L24529; AAA99053.1; -  
CC DR EMBL; L11242; AAA99053.1; JOINED.  
CC DR EMBL; L24502; AAA99053.1; JOINED.  
CC DR EMBL; L24522; AAA99053.1; JOINED.  
CC DR EMBL; L24523; AAA99053.1; JOINED.  
CC DR EMBL; L24525; AAA99053.1; JOINED.  
CC DR EMBL; L24526; AAA99053.1; JOINED.  
CC DR EMBL; L24527; AAA99053.1; JOINED.  
CC DR EMBL; L24528; AAA99053.1; JOINED.  
CC DR EMBL; BC014309; AAHL4309.1; -  
CC DR EMBL; C43357; C43357.  
CC DR PDB; ILPA; 01-NOV-94.  
CC DR PDB; ILPB; 01-NOV-94.  
CC DR PDB; 1GPL; 12-FEB-97.  
CC DR PDB; 1N8S; 18-DEC-02.  
CC DR Genew; HGNC:9155; PNLIP.  
CC DR MM; 246600; -  
CC DR GO; GO:0004806; F:triacylglycerol lipase activity; TAS.  
CC DR InterPro; IPR000734; Lipase.  
CC DR InterPro; IPR001024; Lipoxigenase LH2.  
CC DR InterPro; IPR000379; Ser esters\_site.  
CC DR Pfam; PF00151; lipase; 1.  
CC DR Pfam; PF01477; PLAT; 1.  
CC DR PRINTS; PR00821; TAGLIPASE.  
CC DR SMART; SMC0309; LH2; 1.  
CC DR PROSITE; PS00120; LIPASE\_SER; 1.  
CC DR PROSITE; PS50095; PLAT; 1.  
CC KW Hydrolase; Lipid degradation; Pancreas; Glycoprotein; Signal;  
CC KW 3D-structure.  
CC FT SIGNAL 1 16  
CC FT CHAIN 17 465 TRIACYLGLYCEROL LIPASE, PANCREATIC.  
CC FT DOMAIN 355 465 PLAT.  
CC FT ACT\_SITE 169 169 CHARGE RELAY SYSTEM.  
CC FT ACT\_SITE 193 193 CHARGE RELAY SYSTEM.  
CC FT ACT\_SITE 280 280 CHARGE RELAY SYSTEM.  
CC FT DISULFID 20 26  
CC FT DISULFID 107 118  
CC FT DISULFID 254 278  
CC FT DISULFID 302 313  
CC FT DISULFID 316 321  
CC FT DISULFID 449 465  
CC FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .).  
CC FT STRAND 18 21  
CC FT TURN 22 24  
CC FT TURN 25 28  
CC FT STRAND 31 33  
CC FT TURN 48 51  
CC FT STRAND 54 58  
CC FT STRAND 67 69  
CC FT HELIX 73 78  
CC FT TURN 83 84  
CC FT STRAND 87 91  
CC FT TURN 94 95  
CC FT TURN 98 99  
CC FT HELIX 101 110  
CC FT TURN 111 113  
CC FT STRAND 117 122  
CC FT HELIX 124 127  
CC FT HELIX 131 136  
CC FT HELIX 160 162  
CC FT STRAND 163 168  
CC FT TURN 169 169  
CC FT HELIX 170 180  
CC FT TURN 181 184  
CC FT STRAND 188 192

STRAND 195 195  
TURN 197 201  
TURN 204 206  
HELIX 210 212  
STRAND 216 219  
TURN 227 229  
STRAND 233 233  
STRAND 241 245  
TURN 246 247  
TURN 252 253  
TURN 258 262  
TURN 263 263  
TURN 266 266  
HELIX 268 291  
TURN 292 292  
TURN 294 297  
STRAND 298 298  
STRAND 300 301  
HELIX 305 309  
TURN 310 311  
TURN 318 319  
STRAND 323 323  
TURN 326 327  
TURN 328 330  
TURN 332 335  
STRAND 340 344  
STRAND 355 365  
STRAND 368 378  
TURN 379 380  
STRAND 381 382  
STRAND 386 389  
TURN 395 396  
STRAND 398 405  
STRAND 412 420  
TURN 426 427  
STRAND 431 440  
TURN 441 442  
STRAND 445 449  
STRAND 454 454  
TURN 456 457  
STRAND 460 465  
SEQUENCE 465 AA; 51157 MW; 2BC49CC7F0E2DF52 CRC64;

RY Match 44.1%; Score 1115; DB 1; Length 465;  
Local Similarity 47.8%; Pred. No. 1.3e-84;  
Ches 226; Conservative 73; Mismatches 160; Indels 14; Gaps 10;

1 MLGIVTAPLFGTGRGKVCYERLGCDFKGLPWTRTFSTELVGLPKSPKINTRELLYT 60  
1 MLPLWTL-LLLGAVAGKVCYERLGCSDSPSGITERPLHLPSPKDVNTRFLLYT 59

61 IHNPNAYQISAVNSSTTOASVFGTDKITRINIAGW--KTDGKQWQDMCNVLLQLEDINC 118  
60 NENPNFQEVAA--DSSISGSGNFKTRNRTREIIEGFIDKGEENLANVCNLPKVSVC 118

119 INLDWTNGSRE-YIHAVNLRVGAEVAYFDVLMKKFEPYSPKVLHIGSLGNAAGEA 177  
119 ICVDKXGSGRTCYTOASQNRIRVGAEVAYFEVLFQSAFGYSPSNVHVHIGSLGNAAGEA 178

178 GSRIPG-LGRITGLDPAGFPFFHTNTPKVRLPDSDANFVDVHTHTNAARILFELGVGTIDAC 236  
179 GRRTNGTIGRTGLDPAEPFCQGPPELVRLDPSDAKFVDVHTHTDGAPIVNLFGMSQVV 238

237 GHLDYFPGGKMPCELDLITPLKFNENAYKEMASFFDCNHRASQVFAESILNPDAP 296  
239 GHLDYFPGGKMPCELDLITPLKFNENAYKEMASFFDCNHRASQVFAESILNPDAP 297

297 IAYPCRSYTSFAGNCFCTSCGEGCTMGHFADRHFNKMTN--GSHYFLNTGSLSPPEAR 354  
298 AGFPACASNVFTANKCFPCGGGCPQNGHYADRY---PGKTNVGGQFYLDTDGASNFAR 354

355 WHHKLKSVKLSGSEVTOGTVFLRVGAIGKTGEFAIVSGKLEBPGMTYTKLIDADVNGNIT 414

Db 355 WRYKVSVTLSGKKVT-GHILVSLFGNKGNSKQYEIFKGTLPKDPSTHNSFDSVDVGDQLQ 413  
Qy 415 SVQFIWKXHLFEDSQNKLGAEMVINTSGYKSTFGSQDIMGNINQLNKLKPC 467  
Db 414 MKKFIWYNNVNPITLPRVGASKII-VETNKGQNFCSPEIVREVLTLTTPC 465

RESULT 9  
LIPP\_HORSE  
ID LIPP HORSE STANDARD; PRT; 461 AA.  
AC P29183;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Triacylglycerol lipase, pancreatic precursor (EC 3.1.1.3) (Pancreatic lipase) (PL) (fragment).  
GN PNLIP.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=92267022; PubMed=1587279;  
RA Kerfelec B., Foglizzo E., Bonicel J., Bougis P.E., Chapus C.;  
RT "Sequence of horse pancreatic lipase as determined by protein and  
RT cDNA sequencing. Implications for p-nitrophenyl acetate hydrolysis by  
RT pancreatic lipases.";  
RL Eur. J. Biochem. 206:279-287(1992).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
RX MEDLINE=94238688; PubMed=8182745;  
RA Bourne Y., Martinez C., Kerfelec B., Lombardo D., Chapus C.,  
RA Cambilliau C.;  
RT "Horse pancreatic lipase. The crystal structure refined at 2.3-A  
RT resolution.";  
RL J. Mol. Biol. 238:709-732(1994).  
CC -!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a  
CC fatty acid anion.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.  
CC -!- SIMILARITY: Contains 1 PLAT domain.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; X66218; CAA46961.1; -.  
PIR; S21223; S21223.  
PDB; 1HPL; 31-MAY-94.  
InterPro; IPR000734; Lipase.  
InterPro; IPR001024; Lipoxigenase LH2.  
InterPro; IPR000379; Ser\_estrs\_site.  
Pfam; PF00151; lipase; 1.  
Pfam; PF01477; PLAT; 1.  
PRINTS; PR00821; TAGLIPASE.  
SMART; SM00308; LH2; 1.  
PROSITE; PS00120; LIPASE\_SER; 1.  
PROSITE; PS50095; PLAT; 1.  
Hydrolase; Lipid degradation; Pancreas; Signal; 3D-structure.  
FT NON TER 1 12  
FT SIGNAL <1 12  
FT CHAIN 13 461  
FT DOMAIN 351 461  
FT ACT\_SITE 165 165  
FT ACT\_SITE 189 189  
FT ACT\_SITE 276 276  
FT DISULFID 16 22  
FT DISULFID 103 114

TRIACYLGLYCEROL LIPASE, PANCREATIC.  
PLAT.  
CHARGE RELAY SYSTEM.  
CHARGE RELAY SYSTEM.  
CHARGE RELAY SYSTEM.

DISULFID 250  
DISULFID 298  
DISULFID 312  
DISULFID 445  
MUTAGEN 165  
S->X: 80% LOSS OF P-NITROPHENYL ACETATE  
HYDROLYSIS ACTIVITY.  
14 17  
TURN 18 20  
STRAND 21 24  
TURN 27 28  
STRAND 32 34  
TURN 33 34  
STRAND 35 35  
HELIX 44 47  
STRAND 50 55  
TURN 56 57  
STRAND 58 65  
HELIX 69 74  
TURN 79 80  
STRAND 82 87  
TURN 93 94  
TURN 96 97  
HELIX 109 98  
STRAND 112 118  
HELIX 120 123  
HELIX 127 152  
HELIX 156 158  
STRAND 159 164  
TURN 165 165  
HELIX 166 177  
TURN 178 180  
STRAND 184 188  
STRAND 191 191  
TURN 193 197  
TURN 200 202  
HELIX 206 208  
STRAND 211 215  
HELIX 222 225  
STRAND 229 229  
STRAND 237 241  
TURN 242 243  
TURN 248 249  
HELIX 261 265  
TURN 266 267  
HELIX 274 288  
TURN 290 291  
STRAND 294 294  
STRAND 296 297  
HELIX 301 305  
TURN 306 307  
TURN 314 315  
STRAND 319 319  
TURN 322 323  
HELIX 324 326  
TURN 328 331  
STRAND 336 340  
STRAND 351 361  
STRAND 364 374  
TURN 375 376  
STRAND 377 378  
STRAND 382 389  
STRAND 394 401  
STRAND 408 417  
TURN 422 423  
STRAND 427 435  
TURN 437 438  
STRAND 441 445  
STRAND 450 450  
TURN 452 453  
STRAND 456 461  
SEQUENCE 461 AA; 50921 MW; 382F33F3CE446738 CRC64;  
43.9%; Score 1109; DB 1; Length 461;  
ry Match

Best Local Similarity 48.4%; Pred. No. 3.9e-84;  
Matches 226; Conservative 66; Mismatches 165; Indels 10; Gaps 9;  
QY 5 WIVAFPFGRSRKVCYERLGCYKDPWTFRTFTELVLGFWPEKINTRELLYTHNP 64  
DB 1 WTLS-LLLGAVGVNEVCYERLGCYKDPWTFRTFTELVLGFWPEKINTRELLYTHNP 59  
QY 65 NAYQEI-SAVNSSTIOASYGTDKITRINIAGW--KTDGKQORDMCMVLLQLEDINCINLD 122  
DB 60 DNFEIVA-DPSTIQSSNENTGRKTRFIHGHIDGERSWLSTMCQNMFKVESVNCICVD 118  
QY 123 WINGSR-EYIHAVNNRVVGAEVAYFIDVLMKKFYSKPVHLSLGAHLAGEAGSRI 181  
DB 119 WKSGRATAYSQASQNVRIVGAEVAYLVGLQSPFSPSNVHIIHGLSLGSHAAAGAGRT 178  
QY 182 PG-LGRIITGLDPAGPFPHNTKPEVRLDPSDANFVHTNAARILFELGVGTIDACGHL 240  
DB 179 NGVGRITGLDPAEPCFQGTPELVRLDPSDAQFVDVIHTDIAPFIPNLFQSGMSTAGHL 238  
QY 241 FYPNGGKHPGCCDLITPLLKFNFNAYKKEVAFDCHNARSYQFYAESILNPDATAYP 300  
DB 239 FFPNGGKEMFGCKQNVLSQI-VDIDGIQGTDRDFAACNHLRSYKYVYDSILNPDGAFGS 297  
QY 301 CRSYTSFKAGNCFPCSKGECPTMCHPADRFHFKNMKTNGSHYFLNTGSLSPFARWKL 360  
DB 298 CASYSDFTANKCFPCSGEGCPQMGHYADRFPGRT-KGVGQLFYLTGDSANFARWRV 356  
QY 361 VKLSGSEVTQGTVFLRVGGAIGTGEFAIVSGKLEPGMTYTKLIDADVNVGNITSVOFIW 420  
DB 357 VTLSGKVT-GHVLVSLFGNKSQRQYEIFQGTCLKPDNTYSNEFSDSDEVDGDKVKFIW 415  
QY 421 KXHLFEDSQKLGAEWINTSGKYKSTFCSDQINGNFIQNLKPC 467  
DB 416 YNNVINLTLPKVGASK-ITVERNDGVSFNFCSSETVREDVLLTLTAC 461  
RESULT 10  
LIPP\_CAVPO STANDARD; PRT; 465 AA.  
AC PS0903;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Triacylglycerol lipase, pancreatic precursor (EC 3.1.1.3) (Pancreatic  
lipase) (PL).  
GN PNLIP.  
OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
OX NCBI\_TaxID=10141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pancreas;  
RX MEDLINE=94139898; PubMed=8307159;  
RA Carriere F., Thirstrup K., Hjorth S., Boel E.;  
RT "Cloning of the classical guinea pig pancreatic lipase and comparison  
with the lipase related protein 2";  
RL FEBS Lett. 338:63-68(1994).  
CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a  
fatty acid anion.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.  
CC -1- SIMILARITY: Contains 1 PLAT domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X77403; CAA54585.1; -.



[illegible]

T 11  
RABI  
LIPP  
2021  
01-JU  
01-JU  
5-SI

232 TIDAGHLDVYPNGKMPGCEDLITLLKFNFRAYKKEVASFDCNHRASVQFVBSIL 291  
 235 MSQTVGHLDFPFNGKMPGCKQKVLQSQI-VDINGVWEGTRDFVACNHLRSYKYVADSI 293  
 292 NPDAFIAPYCRSYTSFKAGNCGFFCSKSGCPTMGHFAFRHFKNMKTN--GSHVFLNTGSL 349  
 294 NPNGFAGFSCASYTAFSANKCFPCS-NGCFQMGHYADRF---SRKTDGVGTQFYLTNGDS 349  
 350 SPFAEMHKLKSVKLSGSEVYTGTVFLVGAIGAKTGEFAIVSGKLEPQGYTYKLLDADVN 409  
 350 SNFARWYQVAVLTUSGRVTVTGHVLSLYSGKNSQYEIFTGTLKPGDTHLNEFSDVD 408  
 410 VGNITSVQFVTKHFLFSDSQNKLAENAVINTSGKYGVKSTFCSQDIMGNNILQNLKPC 467  
 409 GVDYQVKVFWYNNVNTPLPKVGASQ-ITVEQNDGRVFKFCSTDVREDILLTLTPC 465

JT 12  
 -MOUSE  
 -LIP2 MOUSE STANDARD; PRT; 468 AA.  
 P17892;  
 01-NOV-1990 (Rel. 16, Created)  
 01-NOV-1990 (Rel. 16, Last sequence update)  
 15-SEP-2003 (Rel. 42, Last annotation update)  
 Pancreatic lipase related protein 2 precursor (BC 3.1.1.3) (Cytotoxic  
 T lymphocyte lipase).  
 PNLIPR2 OR PLRP2.  
 Mus musculus (Mouse).  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxId=10090;  
 [1]  
 SEQUENCE FROM N.A.  
 TISSUE=T-cell;  
 MEDLINE=90150272; PubMed=2302735;  
 Grusby M.J., Nabavi N., Wong H., Dick R.F., Bluestone J.A.,  
 Schotz M.C., Glimcher L.H.  
 "Cloning of an interleukin-4 inducible gene from cytotoxic T  
 lymphocytes and its identification as a lipase."  
 Cell 60:451-459(1990).  
 -!- FUNCTION: MAY HAVE AN IMPORTANT ROLE IN CYTOTOXIC T LYMPHOCYTE  
 (CTL) EFFECTOR FUNCTION.  
 -!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a  
 fatty acid anion.  
 -!- INDUCTION: By interleukin-4.  
 -!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.  
 -!- SIMILARITY: Contains 1 PLAT domain.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcements>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; M30687; AAA37491.1; ALT\_INIT.  
 HSP; P54318; 1BU8.  
 MGD; MGI:1336202; Pnlipr2.  
 GO; GO:0006968; P:cellular defense response; IMP.  
 GO; GO:0007586; P:digestion; IMP.  
 InterPro; IPR000734; Lipase.  
 InterPro; IPR001024; Lipoxigenase LH2.  
 InterPro; IPR000379; Ser esterase site.  
 Pfam; PF00151; lipase; 1.  
 Pfam; PF01477; PLAT; 1.  
 PRINTS; PR00821; TAGLIPASE.  
 SMART; SM00308; LH2; 1.  
 PROSITE; PS00120; LIPASE\_SER; FALSE\_NEG.  
 PROSITE; PS00095; PLAT; 1.  
 Hydrolyase; lipid degradation; Pancreas; Glycoprotein; Signal.  
 SIGNAL 16  
 [2]

FT CHAIN 17 468 PANCREATIC LIPASE RELATED PROTEIN 2.  
 FT DOMAIN 356 468 PLAT.  
 FT ACT\_SITE 170 170 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 194 194 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 281 281 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT DISULFID 20 26 BY SIMILARITY.  
 FT DISULFID 108 119 BY SIMILARITY.  
 FT DISULFID 255 279 BY SIMILARITY.  
 FT DISULFID 303 314 BY SIMILARITY.  
 FT DISULFID 452 468 BY SIMILARITY.  
 FT CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 468 AA; 52585 MW; 1698E35C0P4E1534 CRC64;  
 Query Match 43.3%; Score 1094.5; DB 1; Length 468;  
 Best Local Similarity 46.7%; Pred. No. 6.3e-83;  
 Matches 220; Conservative 68; Mismatches 176; Indels 7; Gaps 6;  
 Qy 1 MLGIMIVAFLEFSGRGKEVCYERLGCFCGDLPTWTRTFSTELVGLPWSPEKINRFLLYT 60  
 Db 1 MLLCWLVS-LLLATVCGKEVCYHGLGCFSDNPKPMAGMIQRPSKIFPWSPEIDITRFLLYT 59  
 Qy 61 IHNPNAYOISAVNSSTICASYFTGDKITRINIAWKTDGK--WORDMNVLLQLEDINC 118  
 Db 60 NENPNYQIISATDPATINASFQDLKTRFTIHFIDKBEGLMDCKKQFQVEKVC 119  
 Qy 119 INLDWINGSR-EYIEHAVNNLRVVGAEVAFIDVLMKPEYSPKVHLIGHSLGAHLAGEA 177  
 Db 120 ICVDWKRGSRTEYQTASYNTRVLGAEIAFLVQLSTENGSPENVHLIPHSLGSHVAGEA 179  
 Qy 178 GSRIPG-LGRIITGLDPAGFFHTPKFVRLDPSDANFVDVHTNAARILFELGVGTIDAC 236  
 Db 180 GRLEGHVGRITGLDPAEPFCQGLPEEVRLDPSDAMFVDVHTDSAPIIPYLGFGMSQKV 239  
 Qy 237 GHLDFYPNGKMPGCEDLITELLKFNFNAYKEMASPFDCNHRASVQFVBSILNDPAF 296  
 Db 240 GHLDFYPNGKMPGCEDLITELLKFNFNAYKEMASPFDCNHRASVQFVBSILNDPAF 298  
 Qy 297 IAYPCSYTSFKAGNCGFFCSKSGCPTMGHFAFRHFKNMKTNGSHYFLNTGSLSPFARWR 356  
 Db 299 LGYPCSSYFKFQHNDCFFSPGECQCPKMGHYADQFEGKT-ATVEQTFELNTADTGNFTWR 357  
 Qy 357 HKLSVKLSSEVYTGTVFLRVGGAIGKTEGPAIVSGKLEPQGYTYKLLDADVNYGNITSV 416  
 Db 358 YKVSVTLSPKSLSGVILVALYGCNANSQKYEVEFKGSLQPEARYIRIDVDVNVGEIQKV 417  
 Qy 417 QFIWKHLPEDSQNKLGAEVINTSGKYGVKSTFCSQDIMGNNILQNLKPC 467  
 Db 418 KFLNNKVINLFRPTMGASQITLQAKDGKBFNCTSNVHEDVLQSLYPC 468

RESULT 13  
 LIPP RAT STANDARD; PRT; 465 AA.  
 ID LIPP RAT AC P27657;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Triacylglycerol lipase, pancreatic precursor (EC 3.1.1.3) (Pancreatic  
 lipase) (PL).  
 GN PNLIP.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxId=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94262798; PubMed=8203536;  
 RA Payne R.M., Sims H.F., Jennens M.L., Lowe M.E.;  
 RT "Rat pancreatic lipase and two related proteins: enzymatic properties  
 and mRNA expression during development.";  
 RL Am. J. Physiol. 266:G914-G921(1994).  
 [2]

SEQUENCE FROM N.A., AND SEQUENCE OF 17-37.  
 STRAIN=Sprague-Dawley; TISSUE=Pancreas;  
 MEDLINE=93252914; PubMed=8486693;  
 Wishart M.J., Andrews P.C., Nichols R., Blevins G.T. Jr.,  
 Logsdon C.D., Williams J.A.,  
 "Identification and cloning of GP-3 from rat pancreatic acinar  
 zymogen granules as a glycosylated membrane-associated lipase";  
 J. Biol. Chem. 268:10303-10311(1993).  
 -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN FAT METABOLISM. IT  
 PREFERENTIALLY SPLITs THE ESTERS OF LONG-CHAIN FATTY ACIDS AT  
 POSITIONS 1 AND 3, PRODUCING MAINLY 2-MONOACYLGLYCEROL AND FREE  
 FATTY ACIDS, AND SHOWS CONSIDERABLY HIGHER ACTIVITY AGAINST  
 INSOLUBLE EMULSIFIED SUBSTRATES THAN AGAINST SOLUBLE ONES.  
 -!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a  
 fatty acid anion.  
 -!- SUBCELLULAR LOCATION: Secreted.  
 -!- TISSUE SPECIFICITY: PANCREAS. PANCREATIC SECRETORY (ZYMOGEN)  
 GRANULE.  
 -!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.  
 -!- SIMILARITY: Contains 1 PLAT domain.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M58369; AA79888.1; ..  
 HSSP; P00591; 1ETH.  
 InterPro: IPR000734; Lipase.  
 InterPro: IPR001024; Lipoxigenase LH2.  
 InterPro: IPR000379; Ser\_estrs\_site.  
 Pfam; PF00151; Lipase; 1.  
 Pfam; PF01477; PLAT; 1.  
 PRINTS; PRC0821; TAGLIPASE.  
 SMART; SM00308; LH2; 1.  
 PROSITE; PS00120; LIPASE SER; 1.  
 PROSITE; PS00095; PLAT; 1.  
 Hydrolase; Lipid degradation; Pancreas; Signal.  
 SIGNAL  
 CHAIN 1 16  
 DOMAIN 17 465 TRIACYLGLYCEROL LIPASE, PANCREATIC.  
 ACT\_SITE 355 465 PLAT.  
 ACT\_SITE 169 169 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 ACT\_SITE 193 193 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 ACT\_SITE 280 280 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 DISULFID 20 26 BY SIMILARITY.  
 DISULFID 107 118 BY SIMILARITY.  
 DISULFID 254 278 BY SIMILARITY.  
 DISULFID 302 313 BY SIMILARITY.  
 DISULFID 316 321 BY SIMILARITY.  
 DISULFID 449 465 BY SIMILARITY.  
 SEQUENCE 465 AA; 51440 MW; CP014D58AA09A6A2 CRC64;

ty Match 42.5%; Score 1073; DB 1; Length 465;  
 - Local similarity 46.1%; Pred. No. 3.7e-81;  
 -hes 217; Conservative 78; Mismatches 166; Indels 10; Gaps 9;  
 1 MGIWVAFLEFGTSRGVCYERLGGCFKGLPWTRTFSTELVCLPSPKINTRFLVLT 60  
 1 MMLMTFAVL-LGAVAGEVCFDKLGGCFSDPAGWSGTRDPLKALPSPAQINTRFLVLT 59  
 61 IHNPNAYGEISAVNSSTIQASYFGDTKTRINIAGW--KTDGKQWDMCNVLQLEDINC 118  
 60 NENQDNQKITS-DASSIRNSFNKTRKTRIIHGFIDKGEENWLSMDCKMFKVESVNC 118  
 119 INLDWINGSR-EYHAVNVLNVEGAEVAFIDVLKMFYSKVLHIGSLGHAUAGA 177  
 119 ICVDWKGSRATYQATQCNVRWGAEVALLVNLKSLDGLGHPDPNVHLIGHSLGSHVAGEA 178  
 178 GSRIPG-LGRITGLDPAGFPFFHPKPEVRLDPSDANFVDVHTNNAIRILFELGVGTIDAC 236

Db 179 GKRTGAIGRITGLDAAEPYFQGTPEEVRLDPTDAQFVDAIHTDAAPILNLFQMSQTV 238  
 QY 237 GHLDFYPNGGKMPCCEDLIITPLKFNFNAYKKEMASFFDCNHFARSYQFYAESILNPDAP 296  
 Db 239 GHLDFPFGNGMEMPGCCQXNLSQI-VDIGIWEGRDFAACNHLRSYKYVTDSIVNPTGF 297  
 QY 297 IAYPCRSTYTSFAGNCFPCSEGGPTWGHFADRFHFKNMKINGSHYFLNTGSLSPFARWR 356  
 Db 298 SGFSCSSYNVFSANKFCFCGSEGGPQMGHYADKYPGKT-KELYOKFVLNTGDKSNFARWR 356  
 QY 357 HKLSVLSGSEVTOGTFLRVGGAIKGTGEPFAIVSGKLEPQMTYTKLIDADVNVGNITSV 416  
 Db 357 YQVTVTLSGQKVT-GHILVSLFNGGNSKQYEVFKGSLHPGDTHVKEFDSMDVGDLOKV 415  
 QY 417 QPINKHLFEDSQNKLGAEMVINTSGKYGKSTFCSDINGPNILQNLKPC 467  
 Db 416 KFIWNVNVTLPKVGASR-ISVERNDGRVFNFCSDQTVREDVILLTSLAC 465

## RESULT 14

LIPP PIG  
 ID LIPP PTG STANDARD; PRT; 450 AA.  
 AC P00591;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Triacylglycerol lipase, pancreatic (EC 3.1.1.3) (pancreatic lipase)  
 DE (PL)  
 GN PNLIP.  
 OS Sus scrofa (Pig)  
 OC Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE OF 308-449.  
 RX MEDLINE=82113655; PubMed=7326260;  
 RA de Caro J.D., Boudouard M., Bonicel J.J., Guidoni A.A., Desnuelle P.,  
 Roverly M.;  
 RA "Porcine pancreatic lipase. Completion of the primary structure.";  
 RT Biochim. Biophys. Acta 671:129-138(1981).  
 RN [2]  
 RP SEQUENCE OF 1-234, AND CARBOHYDRATE-LINKAGE SITE.  
 RX MEDLINE=79236335; PubMed=380992;  
 RA Bianchetta J.D., Bidaud J., Guidoni A.A., Bonicel J.J., Roverly M.;  
 RT "Porcine pancreatic lipase. Sequence of the first 234 amino acids of  
 the peptide chain.";  
 RL Eur. J. Biochem. 97:395-405(1979).  
 RN [3]  
 RP SEQUENCE OF 235-307.  
 RX MEDLINE=8008846; PubMed=518929;  
 RA Guidoni A.A., Bonicel J.J., Bianchetta J.D., Roverly M.;  
 RT "Porcine pancreatic lipase. Sequence between the 235th and 307th  
 amino acids.";  
 RL Biochimie 61:841-845(1979).  
 RN [4]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=83105095; PubMed=7151781;  
 RA Benkouna F., Guidoni A.A., de Caro J.D., Bonicel J.J.,  
 Desnuelle P.A., Roverly M.;  
 RT "Porcine pancreatic lipase. The disulfide bridges and the sulfhydryl  
 groups.";  
 RL Eur. J. Biochem. 128:331-341(1982).  
 RN [5]  
 RP SUBSTRATE-BINDING SITE.  
 RX MEDLINE=82000578; PubMed=6791692;  
 RA Guidoni A.A., Benkouna F., de Caro J.D., Roverly M.;  
 RT "Characterization of the serine reacting with diethyl p-nitrophenyl  
 phosphate in porcine pancreatic lipase.";  
 RL Biochim. Biophys. Acta 660:148-150(1981).  
 RN [6]  
 RP STRUCTURE OF CARBOHYDRATE.  
 RX MEDLINE=88082841; PubMed=3691527;  
 RA Fournet B., Leroy Y., Montreuil J., Decaro J., Roverly M.,

van Kuik J.A., Vliegthart J.F.G.;  
"Primary structure of the glycans of porcine pancreatic lipase.";  
Eur. J. Biochem. 170:369-371(1987).  
[7]  
X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS TO 30-32.  
MEDLINE=98273947; PubMed=8663362;  
Hermoso J., Pignol D., Kerfelec B., Crenon I., Chapus C.,  
Fontecilla-Camps J.C.;  
"Lipase activation by nonionic detergents. The crystal structure of  
the porcine lipase-collipase-tetraethylene glycol monoethyl ether  
complex.";  
J. Biol. Chem. 271:18007-18016(1996).  
-!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a  
fatty acid anion.  
-!- SUBCELLULAR LOCATION: Secreted.  
-!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.  
-!- SIMILARITY: Contains 1 PLAT domain.  
-!- DATABASE: NAME=Worthington enzyme manual;  
WWW="http://www.worthington-biochem.com/manual/L/PL.html".  
PDB; IETH, 07-DEC-96.  
InterPro; IPR000734; Lipase.  
InterPro; IPR001024; lipoxigenase\_LH2.  
InterPro; IPR000379; Ser esters\_site.  
Pfam; PF00151; lipase; 1.  
Pfam; PF01477; PLAT; 1.  
PRINTS; PF00821; TAGLIPASE.  
SMART; SM00308; LH2; 1.  
PROSITE; PS00120; LIPASE\_SER; 1.  
PROSITE; PS00095; PLAT; 1.  
Hydrolase; Lipid degradation; Pancreas; Glycoprotein; 3D-structure.  
DOMAIN 339 450  
ACT\_SITE 153 153 CHARGE RELAY SYSTEM.  
ACT\_SITE 177 177 CHARGE RELAY SYSTEM.  
ACT\_SITE 264 264 CHARGE RELAY SYSTEM.  
DISULFID 4 10  
DISULFID 91 102 IN ISOMER 1.  
DISULFID 91 104 IN ISOMER 2.  
DISULFID 238 262  
DISULFID 286 297  
DISULFID 300 305  
DISULFID 434 450  
CARBOHYD 167 167 N-LINKED (GLCNAC...).  
SEQUENCE 450 AA; 50084 MW; 76E13BDB4541E0E CRC64;  
try Match 42.3%; Score 1069; DB 1; Length 450;  
it Local Similarity 48.7%; Pred.No. 7.7e-81;  
:ches 222; Conservative 62; Mismatches 158; Indels 14; Gaps 10;  
19 EVCYERLCGKDLPTWTRTFSTELVGLPWSPEKINTFLFYTHNPAYQISAVNSSTI 78  
2 EVCPPRLGCPDADPWAGIVQREPLKILPSPKXVDTRFLFYTNQNNYQELVA-DPSTI 60  
79 QASYEGTGDKITRINIAGW--KTGKQWRQMCNVLIQLEPINCINLDWNGRE-YIHAVN 135  
61 TNSNFMDRKTRIFIHGFDKGBEDWLSNCKNLFKVSVCICVDWKGGRGTGTQASQ 120  
136 NLRVGAEVAFYFDVLMKKEFYSPKSVHLIGHSLAGHAGEAGSRIPG-LGRITGLDPAG 194  
121 NIRVGAEVAFYFVELKSSLYGSPSNVHIGSLSHAGEAGRTNGTIERITGLDPAE 180  
195 PFPHNTKPEVRLDPSDANFVDVHTHVAAILFELGVGTIDACGHLDFYNGGKMPGCE 254  
181 PCFGTPELVRLDPSDAKFVDVHTHTDAPIIENLFGMSQTVGHLDFFPNGKQMPGCGK 240  
255 LITPLKFNFNKEMASPFDCNCHARSYQFAESILNPDATFYPCRSYTSFKAGNCFF 314  
241 NILSQI-VDIGIWEGRDPAVACHLSRYKYVADSIINPDGAPGFCDSYVFTANKCF 299  
315 CSKEGCTMGHADRHFHKMKTNG--SHYFNTGSLSPFARWRHKLVLGSGSEVTQGT 372  
300 CPSEGGCPQMGHYADRF---PGKTNQVGSQVFLNTGDSNFAWRVKVSVTLGSKVVT-GH 355  
373 VFLRVGGAIGKTEFAIVSGKLPFGNITYKTLIDADVNTGNTISVQFIW-KXHLFEDSQNK 431

356 ILVSLFGNEGNSQYIEYKGTLPQNTHTSDSDVSDGLQKVFYNNNNVNTPLPR 415  
432 LGAEMVINTSGKYKSTFCSDQIMGPNILQNLKPC 467  
416 VGASK-ITVERNDGKYDFCSQETVREEVLLTNFC 450  
RESULT 15  
LIPH\_HUMAN  
ID LIPH\_HUMAN STANDARD; PRT; 499 AA.  
AC P1150; P78529; Q99465;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Triacylglycerol lipase, hepatic precursor (EC 3.1.1.3) (Hepatic  
lipase) (HL).  
GN LIPC OR HTGL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Liver;  
RX MEDLINE=88273217; PubMed=2839510;  
RA Martin G.A., Busch S.J., Meredith G.D., Cardin A.D., Blankenship D.T.,  
Mao S.J.T., Rehtin A.E., Woods C.W., Racke M.M., Schafer M.P.,  
Fitzgerald M.C., Burke D.M., Flanagan M.A., Jackson R.L.;  
"Isolation and cDNA sequence of human postheparin plasma hepatic  
triglyceride lipase.";  
RL J. Biol. Chem. 263:10907-10914 (1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=88112595; PubMed=2828141;  
RA Stahnke G., Sprengel R., Augustin J., Will H.;  
"Human hepatic triglyceride lipase: cDNA cloning, amino acid sequence  
and expression in a cultured cell line.";  
RL Differentiation 35:45-52 (1987).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=88087233; PubMed=2447084;  
RA Datta S., Luo C.C., Li W.H., Vantuinen P., Ledbetter D.H., Brown M.A.,  
Chen S.H., Liu S., Chan L.;  
"Human hepatic lipase. Cloned cDNA sequence, restriction fragment  
length polymorphisms, chromosomal localization, and evolutionary  
relationships with lipoprotein lipase and pancreatic lipase.";  
RL J. Biol. Chem. 263:1107-1110 (1988).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90105435; PubMed=2805236;  
RA Cai S.J., Wong D.M., Chen S.H., Chan L.;  
"Structure of the human hepatic triglyceride lipase gene.";  
RL Biochemistry 28:8966-8971 (1989).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90216670; PubMed=2324091;  
RA Ameis D., Stahnke G., Kobayashi J., McLean J., Lee G., Buscher M.,  
Schotz M.C., Will H.;  
"Isolation and characterization of the human hepatic lipase gene.";  
RL J. Biol. Chem. 265:6552-6555 (1990).  
RN [6]  
RP SEQUENCE OF 464-499 FROM N.A.  
RX MEDLINE=97018564; PubMed=8865180;  
RA Takagi A., Ikeda Y., Mori A., Ashida Y., Yamamoto A.;  
"Identification of a BstNI polymorphism in exon 9 of the human  
hepatic triglyceride lipase gene.";  
RL Mol. Cell. Probes 10:313-314 (1996).  
RN [7]  
RP VARIANTS PHE-289 AND MET-405.  
RX MEDLINE=93250827; PubMed=1301939;

DR	HSP; P06857; LRPI.	
DR	Genew; HGNC:6619; LIPC.	
DR	MIM; 151670; -	
DR	MIM; 246650; -	
DR	GO; GO:0006487; P-N-linked glycosylation; TAS.	
DR	InterPro; IPR00734; Lipase.	
DR	InterPro; IPR001024; Lipoxigenase_LH2.	
DR	InterPro; IPR000379; Ser_estrse_site.	
DR	Pfam; PF00151; Lipase; 1.	
DR	Pfam; PF01477; PLAT; 1.	
DR	PRINTS; PRO0821; TAGLIPASE.	
DR	SMART; SM00308; LH2; 1.	
DR	PROSITE; PS00120; LIPASE_SER; 1.	
DR	PROSITE; PS00095; PLAT; 1.	
DR	Hydrolase; Lipid degradation; Liver; Heparin-binding; HDL; Signal;	
KX	Polymorphism; Disease mutation.	
KW	SIGNAL	1 22
FT	CHAIN	23 499
FT	DOMAIN	352 486
FT	ACT_SITE	168 168
FT	ACT_SITE	194 194
FT	ACT_SITE	279 279
FT	DOMAIN	191 193
FT	CARBOHYD	42 42
FT	CARBOHYD	78 78
FT	CARBOHYD	362 362
FT	CARBOHYD	397 397
FT	VARIANT	95 95
FT	VARIANT	134 134
FT	VARIANT	215 215
FT	VARIANT	289 289
FT	VARIANT	405 405
FT	VARIANT	440 440
FT	CONFLICT	256 256
FT	CONFLICT	333 333
FT	CONFLICT	499 499
FT	SEQUENCE	499 AA; 55880 MW; A2003B1CAE333C2E CRC64;
QY	Query Match	21.1%; Score 534; dB 1; Length 499;
QY	Best Local Similarity	33.3%; Pred. No. 1.5e-36;
QY	Matches 145; Conservative	64; Mismatches 183; Indels 44; Gaps 14;
QY	51 KINTRELLYTHNPAYQETISAVNSSTIQASFGTDKTRINIAWKTDGK-----WORD 105	
Db	47 ENKTRFLFGTNGOC--QIRNHPDTLQEGFNLSPLVMTIHGWSVGVLENIMQWV 104	
QY	106 MCNVLLQLEDINCINLDWNGSRE-YIHAVNNLRVVGAEVAFIDVLMKKPEYSPKVHL 164	
Db	105 AALKSQAPQVNVGLVDWITLADHYTIARVNTLVGKEVAALLRWLEESVQLSERSVHL 164	
QY	165 IGHSLGAHLAGEAGSRIPG---LGRITGLDPPGPFHNTPEKVRULDPSANFVDVIHTNA 221	
Db	165 IGYSLGARVSGFAGSSIGGTHKTRITGRITDAAAGPLFEQSAENRSLSPDANFVDAIHT-F 223	
QY	222 ARILFELGVGTIDACGHLDPPYNGGKHPGCE--DLITPLKFNFNAYKKEMASFFDCNH 279	
Db	224 TREHWGLSVGIKQPIGHYDFYNGGSGFQGFCHFLRYHIAHQGNA-----ITQTKCSH 279	
QY	280 ARSQVFAEISILNPDA-FTAYPCRSVTSFKAGNCFCKEGCPWTMGHFAFRHFK-NMKT 337	
Db	280 ERSVHLEFIDLSLLHAGTQSMAYPCGMNPSFQGLCLSKCKGRCNTLG-----YHVRQBPSS 334	
QY	338 NGSHYFLNTGSLSPFARWRHKLVLKSLGSEVT--OGTVFLRVGGAIGKTGFALVSGK-L 394	
Db	335 KXKELFLVTRAQSPFKVYHQLKIQINOTETPIQTITMTSLLGTKERQKQIPITLKGKI 394	
QY	395 EPGMTYTKLIDADVNVGNITSVQFIWK-----KHLFEDSQNKLGAEMVINT-- 440	

395 ASNKTYSLITLDVDIGELIMIXFKWENSANVWDTVQTIIPWSTGPRHSGLVLTIR 454

441 --SGKYGYKSTFCSD 454

:|: :|||:

455 VKAGETQORMTFCSEN 470

sh completed: January 6, 2004, 11:23:04  
:ime : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

iclcic - nucleic search, using sw model

m: January 6, 2004, 08:36:36 ; Search time 95 Seconds  
(without alignments)  
6523.178 Million cell updates/sec

ct score: 1404

nce: 1 agcttggatttgattgt.....agaacctgaaccatgtaa 1404

ng table: IDENTITY NUC

Gapop 10.0 , Capext 1.0

hed: 569978 seqs, 220691566 residues

number of hits satisfying chosen parameters: 1139956

um DB seq length: 0

um DB seq length: 2000000000

processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

ase :

- 1: /cgn2.6/prodata/2/ina/5A-COMB.seq.\*
- 2: /cgn2.6/prodata/2/ina/5B-COMB.seq.\*
- 3: /cgn2.6/prodata/2/ina/6A-COMB.seq.\*
- 4: /cgn2.6/prodata/2/ina/6B-COMB.seq.\*
- 5: /cgn2.6/prodata/2/ina/PCUTS-COMB.seq.\*
- 6: /cgn2.6/prodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

t	Score	Match	Length	DB	ID	Description
1	1402.4	99.9	2352	4	US-09-411-132A-1	Sequence 1, Appli
2	1399.4	99.7	1401	4	US-09-411-132A-2	Sequence 2, Appli
3	74.6	5.3	1035	4	US-08-985-492-9	Sequence 9, Appli
4	74.6	5.3	1065	4	US-08-985-492-5	Sequence 5, Appli
5	74.6	5.3	1382	4	US-08-985-492-3	Sequence 3, Appli
6	74.6	5.3	2565	4	US-08-985-492-7	Sequence 7, Appli
7	45.2	3.2	1048	4	US-09-166-205B-63	Sequence 62, Appli
8	43.4	3.1	1341	1	US-08-180-209B-26	Sequence 26, Appli
9	43.4	3.1	1341	1	US-08-385-745-26	Sequence 26, Appli
0	43.4	3.1	1341	3	US-08-485-388-26	Sequence 26, Appli
1	43.4	3.1	1341	3	US-08-474-853-26	Sequence 26, Appli
2	43.4	3.1	1341	4	US-09-166-205B-26	Sequence 26, Appli
3	43.4	3.1	1341	5	PCT-US94-02629-26	Sequence 26, Appli
4	39.2	2.8	8920	2	US-08-446-855A-1	Sequence 1, Appli
5	39.2	2.8	8920	3	US-09-150-741-1	Sequence 1, Appli
6	39.2	2.8	1664976	4	US-08-916-421B-1	Sequence 1, Appli
7	37.2	2.6	231	3	US-08-905-124-4	Sequence 80, Appli
8	37.2	2.6	9734	3	US-09-347-114A-80	Sequence 92, Appli
9	36	2.6	3832	4	US-08-961-527-92	Sequence 102, App
0	35.6	2.5	2341	3	US-08-714-918-102	Sequence 102, App
1	35.6	2.5	2341	3	US-09-265-315-102	Sequence 102, App
2	35.6	2.5	2341	3	US-09-265-315-102	Sequence 102, App
3	35.6	2.5	2341	3	US-09-266-417-102	Sequence 104, App
4	34.8	2.5	1665	4	US-09-601-198-104	Sequence 1, Appli
5	34.6	2.5	26016	4	US-09-326-480A-1	Sequence 1, Appli
6	34.6	2.5	1830121	4	US-09-557-894-1	Sequence 1, Appli
7	34.6	2.5	1830121	4	US-09-643-990A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-411-132A-1  
; Sequence 1, Application US/09411132A  
; Patent No. 6558936

GENERAL INFORMATION:

APPLICANT: Khodadoust, Mehran

APPLICANT: Kapeller-Libermann, Rosana

TITLE OF INVENTION: No. 6558936el Human Lipase Proteins, Nucleic Acides Encoding

TITLE OF INVENTION: Them and uses of Both of these

FILE REFERENCE: 1014-14

CURRENT APPLICATION NUMBER: US/09/411.132A

CURRENT FILING DATE: 2000-09-12

NUMBER OF SEQ ID NOS: 10

SOFTWARE: PatentIn ver. 2.1

SEQ ID NO 1

LENGTH: 2352

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: unsure

LOCATION: (2159)

NAME/KEY: unsure

LOCATION: (2307)

NAME/KEY: unsure

LOCATION: (2313)

US-09-411-132A-1

Query Match 99.9%; Score 1402.4; DB 4; Length 2352;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1403; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ATGCTTGAATTTGGATTGTCATTCTTTGTCATCATCAAGAGGAAAGAGTT	60
DB	125	ATGCTTGAATTTGGATTGTCATTCTTTGTCATCATCAAGAGGAAAGAGTT	184
QY	61	TGCTATGAAGGTTAGGGTGTTCACAGATGTTTACCATGGACCGAGCTTCTCAACA	120
DB	185	TGCTATGAAGGTTAGGGTGTTCACAGATGTTTACCATGGACCGAGCTTCTCAACA	244
QY	121	GAGTTGGTAGGTTTACCTGGTCTCCAGAGAGATAAACACTCGTTTCTGCTTACACT	180
DB	245	GAGTTGGTAGGTTTACCTGGTCTCCAGAGAGATAAACACTCGTTTCTGCTTACACT	304
QY	181	ATACATCCCAATGCCCTATCAGAGATCAGTGGGTTAATTTCAACTATCCAGGCC	240
DB	305	ATACATCCCAATGCCCTATCAGAGATCAGTGGGTTAATTTCAACTATCCAGGCC	364
QY	241	TCATATTTTGAACACAGACAGATCACCGGTATCAACATAGCTGGATGAAACAGATGGC	300
DB	365	TCATATTTTGAACACAGACAGATCACCGGTATCAACATAGCTGGATGAAACAGATGGC	424



301 AAATGGCAGAGAGACATGTGCAATGTGTTCTACAGCTGGAGATATAAATGCAATTAAT 360  
 425 AAATGGCAGAGAGACATGTGCAATGTGTTCTACAGCTGGAGATATAAATGCAATTAAT 484  
 361 TTGATTTGGATCAAGGCTTCAAGGATATCATCCATGCTGTAAACATCTCCGTGTGTT 420  
 485 TTGATTTGGATCAAGGCTTCAAGGATATCATCCATGCTGTAAACATCTCCGTGTGTT 544  
 421 GGTGCTGAGGTGGCTTATTTTATTTGATGTTCTCATGAAAAATTTTGAATATTCCTTCT 480  
 545 GGTGCTGAGGTGGCTTATTTTATTTGATGTTCTCATGAAAAATTTTGAATATTCCTTCT 604  
 481 AAGTGCACCTTGAATGGCCACAGCTTGGAGACACCTGCTGGGAAAGCTGGGTCAAGG 540  
 605 AAGTGCACCTTGAATGGCCACAGCTTGGAGACACCTGCTGGGAAAGCTGGGTCAAGG 664  
 541 ATACAGGCTTGGAGAAATACCTGGGTGGAGCCAGCTGGGCAATTTTCCACAACT 600  
 665 ATACAGGCTTGGAGAAATACCTGGGTGGAGCCAGCTGGGCAATTTTCCACAACT 724  
 601 CCAAGGAGTCAAGCTAGACCTCGATGCGCAACTTTTGTGACGTTATTCATACAAAT 660  
 725 CCAAGGAGTCAAGCTAGACCTCGATGCGCAACTTTTGTGACGTTATTCATACAAAT 784  
 661 GCAGCTCGACCTCTTTGAGCTTGGTGGAGCAATGATGCTGTGCTCATCTTGCAC 720  
 785 GCAGCTCGACCTCTTTGAGCTTGGTGGAGCAATGATGCTGTGCTCATCTTGCAC 844  
 721 TTTTACCAATGAGGAGGAGCAGATCCAGGATGTGAAGACTTAATACACCTTTACTG 780  
 845 TTTTACCAATGAGGAGGAGCAGATCCAGGATGTGAAGACTTAATACACCTTTACTG 904  
 781 AAATTTAACTTCAATGCTTACAAAAGAAATGGCTTCTTTGACGTGAACATGCTC 840  
 905 AAATTTAACTTCAATGCTTACAAAAGAAATGGCTTCTTTGACGTGAACATGCTC 964  
 841 CGAAGTTATCAATTTTATGCTGAAAGCAATCTTAATCTTGAATGCTTATCTTCT 900  
 965 CGAAGTTATCAATTTTATGCTGAAAGCAATCTTAATCTTGAATGCTTATCTTCT 1024  
 901 TGTAGATCTTACACATCTTTTAAAGCAGGAAATGCTTCTTTTGTCCAAAGAGGTTC 960  
 1025 TGTAGATCTTACACATCTTTTAAAGCAGGAAATGCTTCTTTTGTCCAAAGAGGTTC 1084  
 961 CCAACAAATGGGTCAATTTTGTGTGATAGATTTCACTTCAAAAATATGAAGACTTAATGATCA 1020  
 1085 CCAACAAATGGGTCAATTTTGTGTGATAGATTTCACTTCAAAAATATGAAGACTTAATGATCA 1144  
 1021 CATTAATTTTAAACACAGGTCCTTTTCCCATTTTCCCGTTGGAGGCACAAATTTGCT 1080  
 1145 CATTAATTTTAAACACAGGTCCTTTTCCCATTTTCCCGTTGGAGGCACAAATTTGCT 1204  
 1081 GTTAACTCAGTGAAGCAGGTCACCTCAAGGACCTGCTTCTTCTCGTGTAGCGGGCA 1140  
 1205 GTTAACTCAGTGAAGCAGGTCACCTCAAGGACCTGCTTCTTCTCGTGTAGCGGGCA 1264  
 1141 ATTGGGAAAACCTGGGAGTTTGCATTTGTCAGTGGAAAACTTGAGCCAGGATGACTTAC 1200  
 1265 ATTGGGAAAACCTGGGAGTTTGCATTTGTCAGTGGAAAACTTGAGCCAGGATGACTTAC 1324  
 1201 ACAAAATTAATCGATGAGATGTTTAAAGTTGGAAAACTTACAGTGTTCAGTTCATCTGG 1260  
 1325 ACAAAATTAATCGATGAGATGTTTAAAGTTGGAAAACTTACAGTGTTCAGTTCATCTGG 1384  
 1261 AAAAAACATTTGTTTGAAGATTTCTCAGAAATAGTTGGAGCAGAAATGTTGATAATACA 1320  
 1385 AAAAAACATTTGTTTGAAGATTTCTCAGAAATAGTTGGAGCAGAAATGTTGATAATACA 1444  
 1321 TCTGGAAAAATGAGATATAAATCTACCTTCTGTAGCCAGACATTAATGAGGACCTTAATTT 1380  
 1445 TCTGGAAAAATGAGATATAAATCTACCTTCTGTAGCCAGACATTAATGAGGACCTTAATTT 1504

Qy 1381 CTCAGAACCTGAACCACTGCTAA 1404  
 Db 1505 CTCAGAACCTGAACCACTGCTAA 1528  
 RESULT 2  
 US-09-411-132A-2  
 ; Sequence 2, Application US/09411132A  
 ; Patent No. 6558936  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Khodadoust, Mehran  
 ; TITLE OF INVENTION: No. 6558936el Human Lipase Proteins, Nucleic Acids Encoding  
 ; TITLE OF INVENTION: Them, and Uses of Both of These  
 ; FILE REFERENCE: 10147-14  
 ; CURRENT APPLICATION NUMBER: US/09/411,132A  
 ; CURRENT FILING DATE: 2000-09-12  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 1401  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-411-132A-2  
 Query Match 99.7%; Score 1399.4; DB 4; Length 1401;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 ATGCTTGGAAATTTGGATTTGTCATTCTTTTGGCACATCAAGAGGAAAAAGAGTT 60  
 Db 1 ATGCTTGGAAATTTGGATTTGTCATTCTTTTGGCACATCAAGAGGAAAAAGAGTT 60  
 Qy 61 TGTATGAAAGTTAGGTTTCAAGATGTTTACCATGACAGGACCTTCTCAACA 120  
 Db 61 TGTATGAAAGTTAGGTTTCAAGATGTTTACCATGACAGGACCTTCTCAACA 120  
 Qy 121 GAGTTGGTGGTATACCTGGTCTCCAGAGAGATATAACACTCGTTTCTCTACT 180  
 Db 121 GAGTTGGTGGTATACCTGGTCTCCAGAGAGATATAACACTCGTTTCTCTACT 180  
 Qy 181 ATACAAATCCCAATGCCCTATCAGGAGATCAGTGGGTTAAATTTCAACTATCAAGCC 240  
 Db 181 ATACAAATCCCAATGCCCTATCAGGAGATCAGTGGGTTAAATTTCAACTATCAAGCC 240  
 Qy 241 TCATATTTTGGAAACAGACAGATCACCCTGATCAACATAGCTGGATGAAACAGATGGC 300  
 Db 241 TCATATTTTGGAAACAGACAGATCACCCTGATCAACATAGCTGGATGAAACAGATGGC 300  
 Qy 301 AAATGGCAGAGACATGTGCAATGTGTTGTCTACAGCTGGAAGATATAAATGCAATTAAT 360  
 Db 301 AAATGGCAGAGACATGTGCAATGTGTTGTCTACAGCTGGAAGATATAAATGCAATTAAT 360  
 Qy 361 TTGATTTGGATCAAGGTTTCAAGGATATCATCCATGCTGTAAACATCTCCGTGTGTT 420  
 Db 361 TTGATTTGGATCAAGGTTTCAAGGATATCATCCATGCTGTAAACATCTCCGTGTGTT 420  
 Qy 421 GGTGCTGAGGTGGCTTATTTTATTTGATGTTCTCATGAAAAATTTTGAATATTCCTTCT 480  
 Db 421 GGTGCTGAGGTGGCTTATTTTATTTGATGTTCTCATGAAAAATTTTGAATATTCCTTCT 480  
 Qy 481 AAGTGCACCTTGAATGGCCACAGCTTGGAGACACCTGGCTGGGAAAGCTGGGTCAAGG 540  
 Db 481 AAGTGCACCTTGAATGGCCACAGCTTGGAGACACCTGGCTGGGAAAGCTGGGTCAAGG 540  
 Qy 541 ATACAGGCTTGGAGAAATACCTGGGTGGAGCCAGCTGGGCAATTTTCCACAACT 600  
 Db 541 ATACAGGCTTGGAGAAATACCTGGGTGGAGCCAGCTGGGCAATTTTCCACAACT 600  
 Qy 601 CCAAGGAGTCAAGCTAGACCTCGATGCGCAACTTTTGTGACGTTATTCATACAAAT 660  
 Db 601 CCAAGGAGTCAAGCTAGACCTCGATGCGCAACTTTTGTGACGTTATTCATACAAAT 660

661 GCAGCTCGCATCTCTTTGAGCTTGGTGTGGAAACCAATGATGCTTGGTGCATCTTGAC 720  
|||||  
661 GCAGCTCGCATCTCTTTGAGCTTGGTGTGGAAACCAATGATGCTTGGTGCATCTTGAC 720  
|||||  
721 TTTTACCCAAATGGAGGGAAGACATGCGAGATGTGAAGACTTTAAATACACCTTTACTG 780  
|||||  
721 TTTTACCCAAATGGAGGGAAGACATGCGAGATGTGAAGACTTTAAATACACCTTTACTG 780  
|||||  
781 AAATTTAACTTCAATGCTTTACAAAAGAAATGGCTTCTTTCTTTGACTGTAAACCATGCC 840  
|||||  
781 AAATTTAACTTCAATGCTTTACAAAAGAAATGGCTTCTTTCTTTGACTGTAAACCATGCC 840  
|||||  
841 CGAAGTATCAATTTTATGCTGAAGACATCTTAACTCTGATGATTTATGCTTATCTT 900  
|||||  
841 CGAAGTATCAATTTTATGCTGAAGACATCTTAACTCTGATGATTTATGCTTATCTT 900  
|||||  
901 TGTAGATCTACACATCTTTTAAAGCAGGAAATGCTTCTTTCTTCTTCAAGAGGTTGC 960  
|||||  
901 TGTAGATCTACACATCTTTTAAAGCAGGAAATGCTTCTTTCTTCTTCAAGAGGTTGC 960  
|||||  
961 CCAACAAATGGTCAATTTTCTGATGATTTTCACTTCAAAAATATGAAGACTTAATGATCA 1020  
|||||  
961 CCAACAAATGGTCAATTTTCTGATGATTTTCACTTCAAAAATATGAAGACTTAATGATCA 1020  
|||||  
1021 CATTATTTTAAACACAGGGTCCCTTTCCCATTTGCGGTTGGAGGCACAAATGTCT 1080  
|||||  
1021 CATTATTTTAAACACAGGGTCCCTTTCCCATTTGCGGTTGGAGGCACAAATGTCT 1080  
|||||  
1081 GTTAACTCAGTGGAGGAAAGTCACTCAAGAACTGTCTTTCTTCTGTTAGCGGGGCA 1140  
|||||  
1081 GTTAACTCAGTGGAGGAAAGTCACTCAAGAACTGTCTTTCTTCTGTTAGCGGGGCA 1140  
|||||  
1141 ATTGGAAACTGGGGAGTTTGGCAATGTGATGGAACCTTGAGCCAGGATGACTTAC 1200  
|||||  
1141 ATTGGAAACTGGGGAGTTTGGCAATGTGATGGAACCTTGAGCCAGGATGACTTAC 1200  
|||||  
1201 ACAAAATTAATGATGCAAGTGTAAACGTTGGAAACATTAACAAGTGTTCAGTTCATCTGG 1260  
|||||  
1201 ACAAAATTAATGATGCAAGTGTAAACGTTGGAAACATTAACAAGTGTTCAGTTCATCTGG 1260  
|||||  
1261 AAAAAACATTTGTTGAAGATCTCAGATTAAGTTGGAGCAGAAATGTTGATTAATAACA 1320  
|||||  
1261 AAAAAACATTTGTTGAAGATCTCAGATTAAGTTGGAGCAGAAATGTTGATTAATAACA 1320  
|||||  
1321 TCTGGAAATATGGATATAAATCTACCTTCTGTAGCCAAAGACATTAATGGGACCTAATATT 1380  
|||||  
1321 TCTGGAAATATGGATATAAATCTACCTTCTGTAGCCAAAGACATTAATGGGACCTAATATT 1380  
|||||  
1381 CTCCAGAACCTGAAACCATGC 1401  
|||||  
1381 CTCCAGAACCTGAAACCATGC 1401  
|||||

985-492-9  
ence 9, Application US/08985492  
nt No. 6395530

## GENERAL INFORMATION:

APPLICANT: Jaye, Michael C.  
APPLICANT: Doan, Kim-Anh T.  
APPLICANT: Krawiec, John A.  
APPLICANT: Lynch, Kevin J.  
APPLICANT: Amin, Dilip V.  
APPLICANT: South, Victoria J.

TITLE OF INVENTION: LIQ POLYPEPTIDES OF THE TRIACYLGLYCEROL

TITLE OF INVENTION: LIPASE FAMILY, AND COMPOSITIONS AND METHODS FOR THEIR USE

TITLE OF INVENTION: IN ENZYMAIC HYDROLYSIS, AND PROTEIN AND GENE THERAPIES

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Rd. 3C43  
CITY: Collegeville  
STATE: PA

COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/985,492  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Fehner Ph.D., Paul F.  
REGISTRATION NUMBER: 35,135  
REFERENCE/DOCKET NUMBER: A2582-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610)454-3839  
TELEFAX: (610)454-3808  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1035 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1035  
US-08-985-492-9

Query Match 5.3%; Score 74.6; DB 4; Length 1035;  
Best Local Similarity 53.9%; Pred. No. 4e-13;  
Matches 202; Conservative 0; Mismatches 164; Indels 9; Gaps 2;  
QY 388 TACATCCATGCTGTAAACAATCTCCGTTGTTGTTGGTCTGAGGTGGCTTATTTATTGAT 447  
|||  
DB 391 TACACGGATGGGTCAATAATACAGGGTGGTGGACACAGCATTGCCAGGATGCTCGAC 450  
|||  
QY 448 GTTCTCATGAAAAATTTGAATATTCCTTCTTAAAGTGCATTTGATTTGGCCACAGCTTG 507  
|||  
DB 451 TGGCTGACGAGGAGGACGATTTTCTCTCGGAAATGTCCACTTGTATCGGCTACAGCTTC 510  
|||  
QY 508 GGACACACCTGGCTGGGGAAGCTGG---GTCAAGATACACAGGCTTGGAAAGATAACT 564  
|||  
DB 511 GGACGACAGTGGCCGGGTATGAGCACTTCTGTAAGGACGGTGGCCGGAATCACA 570  
|||  
QY 565 GGGTTGACACCTGGCTGGCCATTTTTCACACACTTCCAAAGGAAGTCAAGGTAGACCCC 624  
|||  
DB 571 GGTGATCTCTCGCGGCCCATGTTTGAAGGGCGGACATCCACAAGAGGCTCTCTCGG 630  
|||  
QY 625 TCGGATGCCAACTTTGTTGACGTTATTCACAAATGACAGCTCGCATCTCTTGAGCTT 684  
|||  
DB 631 GACATGACATTTTGGATGTCCTCCACACCTACACGGTTCTCTTCGGCTTGAGCAT 690  
|||  
QY 685 GGTGTTGGAACCATTTGATGCTTGTGTCATTTTACCAATTTTACCCAAATGGAGGAGCAC 744  
|||  
DB 691 GGTATTCAGA-----TGCCTGTGGGCCACATTGACATCTACCCCAATGGGGTGTACTTC 744  
|||  
QY 745 ATGCCAGGATGTGAA 759  
|||  
DB 745 CAGCCAGGCTGTGGA 759  
|||

## RESULT 4

US-08-985-492-5  
; Sequence 5, Application US/08985492  
; Patent No. 6395530

## GENERAL INFORMATION:

APPLICANT: Jaye, Michael C.  
APPLICANT: Doan, Kim-Anh T.  
APPLICANT: Krawiec, John A.  
APPLICANT: Lynch, Kevin J.  
APPLICANT: Amin, Dilip V.

APPLICANT: South, Victoria J.  
 TITLE OF INVENTION: LLG POLYPEPTIDES OF THE TRIACYLGLYCEROL  
 TITLE OF INVENTION: LIPASE FAMILY, AND COMPOSITIONS AND METHODS FOR THEIR USE  
 TITLE OF INVENTION: IN ENZYMAIC HYDROLYSIS, AND PROTEIN AND GENE THERAPIES  
 NUMBER OF SEQUENCES: 31  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Rhone-Poulenc Rorer Inc.  
 STREET: 500 Arcola Rd. 3c43  
 CITY: Collegeville  
 STATE: PA  
 COUNTRY: USA  
 ZIP: 19426  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/985,492  
 FILING DATE:  
 CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fehner Ph.D., Paul F.  
 REGISTRATION NUMBER: 35,135  
 REFERENCE/DOCKET NUMBER: A2582-US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (610)454-3839  
 TELEFAX: (610)454-3808  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1065 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..1065  
 8-985-492-5

ery Match 5.3%; Score 74.6; DB 4; Length 1065;  
 st Local Similarity 53.9%; Pred. No. 4.1e-13;  
 tches 202; Conservative 0; Mismatches 164; Indels 9; Gaps 2;  
 388 TACATCCATGCTGTAACAACTCCCGTGTGTTGGTGTCTGAGGTGCTTATTTATGAT 447  
 391 TACACGGATCGGTCATATATACAGGGTGTGGACACAGCAATGCCAGGATGCTCGAC 450  
 448 GTTCTCATGAAAAAATTGAAATATCCCTTCTTAAAGTGCACTTGATTTGCCACAGCTTG 507  
 451 TGGCTGCAGGAGAGGACGATTTTCTCTCGGAATGTCCACTTGATCGCTACAGCCTC 510  
 508 GGAGCACACCTGGCTGGGAGCTGG--GTCAGGATACCGGCTTGGAGATTAAGTACT 564  
 511 GGAGCGCACGTGGCGGGGTATGAGGCAATCTCGTGAAGGAACGCTGGCGCGAATCACA 570  
 565 GGGTTGGACCCAGCTGGGCGCCATTTTCCACAACTCCAAAGGAAGTCAAGGTAGACCCC 624  
 571 GGTTTGGATCTCGCGGCGCCATGTTTGAAGGCGCGACATCCACAGAGGCTCTCTCG 630  
 625 TCGGATGCCAATTTTGTGAGCTTATTCATAAATGCAAGTCCGATCTCTTTGAGCTT 684  
 631 GACGATGCAGATTTTGTGAGTGTCTCCACCACTTACACGCGTTCCTTCGGCTTGAGCAT 690  
 685 GGTGTGGAACCAATTGATGCTTGTGCTCATCTTGACTTTTACCCAAATGGAGGAGCAC 744  
 691 GGTATTCAGA-----TGCCGTGGGCGACATTTGACATCTACCCCAATGGGGGTGACTTC 744  
 745 ATGCCAGATGTGAA 759  
 745 CAGCCAGGCTGTGA 759

RESULT 5  
 US-08-985-492-3  
 ; Sequence 3, Application US/08985492  
 ; Patent No. 6395530  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jaye, Michael C.  
 ; APPLICANT: Doan, Kim-Anh T.  
 ; APPLICANT: Krawiec, John A.  
 ; APPLICANT: Lynch, Kevin J.  
 ; APPLICANT: Amin, Dilip V.  
 ; APPLICANT: South, Victoria J.  
 ; TITLE OF INVENTION: LLG POLYPEPTIDES OF THE TRIACYLGLYCEROL  
 ; TITLE OF INVENTION: LIPASE FAMILY, AND COMPOSITIONS AND METHODS FOR THEIR USE  
 ; TITLE OF INVENTION: IN ENZYMAIC HYDROLYSIS, AND PROTEIN AND GENE THERAPIES  
 ; NUMBER OF SEQUENCES: 31  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
 ; STREET: 500 Arcola Rd. 3c43  
 ; CITY: Collegeville  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19426  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/985,492  
 ; FILING DATE:  
 ; CLASSIFICATION: 800  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Fehner Ph.D., Paul F.  
 ; REGISTRATION NUMBER: 35,135  
 ; REFERENCE/DOCKET NUMBER: A2582-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (610)454-3839  
 ; TELEFAX: (610)454-3808  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1382 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 312..1370  
 ; US-08-985-492-3  
 Query Match 5.3%; Score 74.6; DB 4; Length 1382;  
 Best Local Similarity 53.9%; Pred. No. 4.7e-13;  
 Matches 202; Conservative 0; Mismatches 164; Indels 9; Gaps 2;  
 Qy 388 TACATCCATGCTGTAACAACTCCCGTGTGTTGGTGTCTGAGGTGCTTATTTATGAT 447  
 Db 702 TACACGGATCGGTCATATATACAGGGTGTGGACACAGCAATGCCAGGATGCTCGAC 761  
 Qy 448 GTTCTCATGAAAAAATTGAAATATCCCTTCTTAAAGTGCACTTGATTTGCCACAGCTTG 507  
 Db 762 TGGCTGCAGGAGAGGACGATTTTCTCTCGGAATGTCCACTTGATCGCTACAGCCTC 821  
 Qy 508 GGAGCACACCTGGCTGGGAGCTGG--GTCAGGATACCGGCTTGGAGATTAAGTACT 564  
 Db 822 GGAGCGCACGTGGCGGGGTATGAGGCGCGACATCCACAGAGGCTCTCTCG 881  
 Qy 565 GGGTTGGACCCAGCTGGGCGCCATTTTCCACAACTCCAAAGGAAGTCAAGGTAGACCCC 624  
 Db 882 GGTTTGGATCTCGCGGCGCCATGTTTGAAGGCGCGACATCCACAGAGGCTCTCTCG 941  
 Qy 625 TCGGATGCCAATTTTGTGAGCTTATTCATAAATGCAAGTCCGATCTCTTTGAGCTT 684  
 Db 942 GACGATGCAGATTTTGTGAGTGTCTCCACCACTTACACGCGTTCCTTCGGCTTGAGCAT 1001

685 GGTGTTGGAACCATGATGCTTGGTCATCTTGATCTTTTACCAATGAGGGAAGCAC 744  
|||||  
1002 GGTATTCAGA-----TGCCTGTGGGCCCAATGACATCTACCCCAATGGGGGTGACTTC 1055  
|||||  
745 ATGCCAGGATGTGAA 759  
1056 CAGCCAGGCTGTGGA 1070  
|||||

T 6

US-985-492-7  
Sequence 7, Application US/08985492  
Patent No. 6395530  
GENERAL INFORMATION:  
APPLICANT: Jaye, Michael C.  
APPLICANT: Doan, Kim-Anh T.  
APPLICANT: Krawiec, John A.  
APPLICANT: Lynch, Kevin J.  
APPLICANT: Amin, Dilip V.  
APPLICANT: South, Victoria J.  
TITLE OF INVENTION: LLG POLYPEPTIDES OF THE TRIACYLGLYCEROL  
LIPASE FAMILY, AND COMPOSITIONS AND METHODS FOR THEIR USE  
TITLE OF INVENTION: LIPASE FAMILY, AND COMPOSITIONS AND METHODS FOR THEIR USE  
TITLE OF INVENTION: IN ENZYMIC HYDROLYSIS, AND PROTEIN AND GENE THERAPIES  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Rd. 3043  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/985,492  
FILING DATE:

CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Fehner Ph.D., Paul F.  
REGISTRATION NUMBER: 35,135  
REFERENCE/DOCKET NUMBER: A2582-US  
TELEPHONE: (610)454-3839  
TELEFAX: (610)454-3808  
FORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2565 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 252...1754  
US-985-492-7

Query Match 5.3%; Score 74.6; DB 4; Length 2565;  
Best Local Similarity 53.9%; Pred. No. 6.8e-13;  
Matches 202; Conservative 0; Mismatches 164; Indels 9; Gaps 2;

388 TACATCCATGCTGTAACAACTCCGCTGTTGTTGCTGAGTGGCTTATTTATTGAT 447  
|||||  
642 TACACGATGCGGTCAATAATACCAGGTGTGGGACACAGCATTGCCAGGATGCTGAC 701  
|||||  
448 GTTCTCATGAAAAAATTTGAATATCCCTTTCTAAAGTGCACTTGGTGGCCACAGCTTG 507  
|||||  
702 TGGCTGCAGGAGGACGATTTTCTCTCGGATGTCCACTTGATCGGCTACAGCCTC 761  
|||||  
508 GGAGCACACCTGGCTGGGGAAGCTGG---GTCAAGGATACAGGCCTTGAAGAATAACT 564  
|||||

Db 762 GGAGCCACGTCGCCGGGTATGCAAGCAACTTCGTGAAGGAACGGTGGCGCAATCACA 821  
|||||  
QY 565 GGGTTGGACCCAGCTGGGCGCATTTTTCACAACACTCCAAAGGAAGTCAGGCTAGACCCC 624  
|||||  
Db 822 GGTTTGGATCCTTGGCGGGCCCATGTTTGAAGGGGCGACATCCACAAGAGCTCTCTCCG 881  
|||||  
QY 625 TCGATGCCAACTTTGTTGACGTTTATTCATAAATGAGCTCGCATCCCTTTTGGAGCTT 684  
|||||  
Db 882 GACGATGCAGATTTTGTGATGTCTCCACACCTACACGCGTTCCTTCGGCTTGAGCAT 941  
|||||  
QY 685 GGTGTTGGAAACCAATGATGCTGTGTCATCTTGACTTTTACCCAAATGGAGGGAAGCAC 744  
|||||  
Db 942 GGTATTCAGA-----TGCCTGTGGGCCCAATGACATCTACCCCAATGGGGGTGACTTC 995  
|||||  
QY 745 ATGCCAGGATGTGAA 759  
|||||  
Db 996 CAGCCAGGCTGTGGA 1010  
|||||

RESULT 7

US-09-166-205B-63  
Sequence 63, Application US/09166205B  
Patent No. 6372471  
GENERAL INFORMATION:  
APPLICANT: Te Piao KING  
TITLE OF INVENTION: CLONING AND RECOMBINANT PRODUCTION OF VESPID VENOM ENZYMES,  
TITLE OF INVENTION: SUCH AS PHOSPHOLIPASE AND HYALURONIDASE, AND IMMUNOLOGICAL  
TITLE OF INVENTION: THERAPIES BASED THEREON  
FILE REFERENCE: 2313/0F138US  
CURRENT APPLICATION NUMBER: US/09/166,205B  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 63  
LENGTH: 1048  
TYPE: DNA  
ORGANISM: Polistes annularis  
US-09-166-205B-63

Query Match 3.2%; Score 45.2; DB 4; Length 1048;  
Best Local Similarity 49.4%; Pred. No. 0.00059;  
Matches 159; Conservative 0; Mismatches 148; Indels 15; Gaps 1;

QY 351 TTGCATTAATTTAGATTGGATCAACGGTTCACGGGAATACATCCATCTGTAAACAATCT 410  
|||||  
Db 318 TTGTAATGCTTTTGCTTCAACAAGGATGCTTTGGGTTATTCAAAGCCGTTGGAACAC 377  
|||||  
QY 411 CCGTGTGTTGCTGCTGAGTGGCTTATTTTATTTGATGTTCTCATGAAAAAATTTGAATA 470  
|||||  
Db 378 ACGTACCGTTGGAAAAATTTGTAGCTGATTTTACAAAACACTACTTGTAGAAAAATATAAAGT 437  
|||||  
QY 471 TTCCCTTCTAAAGTGCACTTGTATGGCCACAGCTTGGGAGCACACTGGCTGGGAAGC 530  
|||||  
Db 438 GCTGATATCAATATACGATGATCGGGCATGTTTGGGCGCCACTACTTCAGGTTTTC 497  
|||||  
QY 531 TGG-----GTCAAGGATACCAGGCTTTGGAAGATAACTGGGTGGACCC 575  
|||||  
Db 498 GGGAAAAAGAGTTCAAAAAGTTAAAAATTAGGAAAAATACAAGAAATTTATCGGCTTGATCC 557  
|||||  
QY 576 AGCTGGGCGCATTTTTCACAACACTCCAAAGGAAGTCAGCTAGACCCCTCGGATGCCAA 635  
|||||  
Db 558 TGCTGGACCGTATTTTTCATCGGAGTGACTGTCCGGACAGACTTTGCGTAACAGACGAGA 617  
|||||  
QY 636 CTTTGTGAGCTTATTCATACA 657  
|||||  
Db 618 ATATGTTCAAGTTATACATACA 639  
|||||

RESULT 8

US-08-180-209B-26  
Sequence 26, Application US/08180209B  
Patent No. 5593877

## GENERAL INFORMATION:

APPLICANT: King, Te-piao  
TITLE OF INVENTION: CLONING AND RECOMBINANT PRODUCTION OF  
VESPID VENOM ENZYMES, SUCH AS PHOSPHOLIPASE AND  
HYALURONIDASE, AND IMMUNOLOGICAL THERAPIES BASED  
THEREON  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klausner & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/180,209B  
FILING DATE: 11-JAN-1994

## CLASSIFICATION:

CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/031,400

## FILING DATE:

FILING DATE: 11-MAR-1993

## ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-074 CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

## INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 1341 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 153..1052

8-180-209B-26

Query Match 3.1%; Score 43.4; DB 1; Length 1341;  
Best Local Similarity 50.9%; Pred. No. 0.0025;  
Matches 141; Conservative 0; Mismatches 121; Indels 15; Gaps 1;

396 TGCTGTAAACAATCTCCGTTGTTGGTGTGCTGAGGTGCTTATTTTATTGATGTTCTCAT 455  
455 TGCTGTAGAAATACACGTTTAGTTGGACAATATATCGCTACGATTAACCCAGAACTCGT 514  
456 GAAAAAATTTGAATATTCCTTCTTAAAGTGCATTTGAGCCACAGCTTGGGAGCACA 515  
515 AAAACACTATAAATCTCGATGCGCAATATACGATTAATTTGGACATAGCTTAGGAGCACA 574  
516 CCTGGCTGGGGAAGCTGGTCAAGGATACAGGCTTGGAG-----AAT 560  
575 TGCTTCAGTTTTTCAGGCAAAAGGTTCAAGATTAAATTTAGGAAATATTTCTGAAT 634  
561 AACTGGGTGGACCCAGCTGGGCCATTTTTCCACAACTCCAAAGGAAGTCAGGCTAGA 620  
635 TATTGGGCTTGATCTGCTAGGCCCTTCGTTTCAATCATTTGTCGGAAGAACTCTG 694  
621 CCCCTCGATGCCAACTTTGTTGACGTTTATTCATACA 657  
695 CGAGACAGATGCAGAAATATGTTCAAATATATACATACA 731

LT 9

3-385-745-26

; Sequence 26, Application US/08385745  
; Patent No. 5612209

## GENERAL INFORMATION:

APPLICANT: King, Te-piao  
TITLE OF INVENTION: Cloning and Recombinant Production of  
Vespid Venom Phospholipases, and Immunological Therapies  
Thereon  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711

## COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/385,745  
FILING DATE:

## CLASSIFICATION:

CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/031,400

## FILING DATE:

FILING DATE:

## ATTORNEY/AGENT INFORMATION:

NAME: Mirock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 3288-020

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 790-9090

TELEFAX: 212 869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 1341 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 153..1052

US-08-385-745-26

Query Match 3.1%; Score 43.4; DB 1; Length 1341;  
Best Local Similarity 50.9%; Pred. No. 0.0025;  
Matches 141; Conservative 0; Mismatches 121; Indels 15; Gaps 1;

Qy 396 TGCTGTAAACAATCTCCGTTGTTGGTGTGCTGAGGTGCTTATTTTATTGATGTTCTCAT 455  
Db 455 TGCTGTAGAAATACACGTTTAGTTGGACAATATATCGCTACGATTAACCCAGAACTCGT 514  
Qy 456 GAAAAAATTTGAATATTCCTTCTTAAAGTGCATTTGAGCCACAGCTTGGGAGCACA 515  
Db 515 AAAACACTATAAATCTCGATGCGCAATATACGATTAATTTGGACATAGCTTAGGAGCACA 574  
Qy 516 CCTGGCTGGGGAAGCTGGTCAAGGATACAGGCTTGGAG-----AAT 560  
Db 575 TGCTTCAGTTTTTCAGGCAAAAGGTTCAAGATTAAATTTAGGAAATATTTCTGAAT 634  
Qy 561 AACTGGGTGGACCCAGCTGGGCCATTTTTCCACAACTCCAAAGGAAGTCAGGCTAGA 620  
Db 635 TATTGGGCTTGATCTGCTAGGCCCTTCGTTTCAATCATTTGTCGGAAGAACTCTG 694  
Qy 621 CCCCTCGATGCCAACTTTGTTGACGTTTATTCATACA 657  
Db 695 CGAGACAGATGCAGAAATATGTTCAAATATATACATACA 731

RESULT 10

-485-388-26  
Sequence 26, Application US/08485388  
ent No. 6270763  
GENERAL INFORMATION:  
APPLICANT: King, Te Piao  
TITLE OF INVENTION: Cloning and Recombinant Production of  
TITLE OF INVENTION: Vespid Venom Phospholipases, and Immunological Therapies  
TITLE OF INVENTION: Based Thereon  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,388  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/385,745  
FILING DATE: 08-FEB-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/031,400  
FILING DATE: 11-MAR-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-074 FMCA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
FORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1341 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 153..1052  
-485-388-26  
Query Match 3.1%; Score 43.4; DB 3; Length 1341;  
t Local Similarity 50.9%; Pred. No. 0.0025;  
ches 141; Conservative 0; Mismatches 121; Indels 15; Gaps 1;  
396 TCGTGTAAACAATCTCCGTTGTGTGGTGGCTGCTTATTTTATTGATGTTCTCAT 455  
455 TCGTGTAGAAATACACGTTTAGTTGGACAATATATCGTACGATTACCCAGAACTCGT 514  
456 GAAAAAATTGAATATTCCTTCTTAAAGTGCACCTTGAATGGCCACAGCTTGGAGCACA 515  
515 AAACACTATAAAATCTCGATGGCAAAATATACGATTAAATTTGGACATAGCTTAGGAGCACA 574  
516 CTGGCTGGGGAAGCTGGGTCAAGGATACAGGCTTGGGAAG-----AAT 560  
575 TCGTTCAGGTTTTCAGGCAAAAAGGTTCAAGAGTTAAATAGGAATAATTTCTGAAT 634  
561 AACTGGTTTGGACCCAGCTGGGCCATTTTCCAACTCCAAAGGAAGTCAAGGCTAGA 620  
635 TATTGGGCTTGATCCTGCTAGGCTTCGTTGCTGATTCATTAATCAATTTCCGAAAGACTCTG 694

QY 621 CCCTCGGATGCCAACTTTTGTGTGACGTTATTATACATA 657  
DB 695 CGAGACAGATGCAGATATGTTCAAATTTATACATA 731  
RESULT 11  
US-08-474-853-26  
Sequence 26, Application US/08474853  
Patent No. 6287559  
GENERAL INFORMATION:  
APPLICANT: King, Te-Piao  
TITLE OF INVENTION: CLONING AND RECOMBINANT PRODUCTION OF  
TITLE OF INVENTION: VESPID VENOM ENZYMES, SUCH AS PHOSPHOLIPASE AND  
TITLE OF INVENTION: HYALURONIDASE, AND IMMUNOLOGICAL THERAPIES BASED THEREON  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,853  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/180,209  
FILING DATE: 11-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/031,400  
FILING DATE: 11-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-074 CIPB  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1341 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 153..1052  
US-08-474-853-26  
Query Match 3.1%; Score 43.4; DB 3; Length 1341;  
Best Local Similarity 50.9%; Pred. No. 0.0025;  
Matches 141; Conservative 0; Mismatches 121; Indels 15; Gaps 1;  
QY 396 TCGTGTAAACAATCTCCGTTGTGTGTGGTGGCTTATTTTATTGATGTTCTCAT 455  
DB 455 TCGTGTAGAAATACACGTTTAGTTGGACAATATATCGCTACGATTACCCAGAACTCGT 514  
QY 456 GAAAAAATTGAATATTCCTTCTTAAAGTGCACCTTGAATGGCCACAGCTTGGAGCACA 515  
DB 515 AAACACTATAAAATCTCGATGGCAAAATATACGATTAAATTTGGACATAGCTTAGGAGCACA 574  
QY 516 CTGGCTGGGGAAGCTGGGTCAAGGATACAGGCTTGGGAAG-----AAT 560  
DB 575 TCGTTCAGGTTTTCAGGCAAAAAGGTTCAAGAGTTAAATAGGAATAATTTCTGAAT 634

561 AACTGGTTGGACCCAGCTGGCCATTTTCCACACACTCCAAAGAGTCAAGCTAGA 620  
562 TTTTGGCTTGGTCTGCTAGGCTTGGTTCGATTCAATTCATTTCCGAAGACTCTG 694  
621 CCCCTCGATGCCAACTTTGTTGACGTTATTCATACA 657  
695 CGAGACAGATGCAGATATGTTCAATTTATACATACA 731

LT 12  
9-166-205B-26  
Sequence 26, Application US/09166205B  
Ent No. 6372471  
GENERAL INFORMATION:  
APPLICANT: Te Piao King  
TITLE OF INVENTION: CLONING AND RECOMBINANT PRODUCTION OF VESPID VENOM ENZYMES,  
TITLE OF INVENTION: SUCH AS PHOSPHOLIPASE AND HYALURONIDASE, AND IMMUNOLOGICAL  
TITLE OF INVENTION: THERAPIES BASED THEREON  
FILE REFERENCE: 2313/OF138US  
CURRENT APPLICATION NUMBER: US/09166.205B  
CURRENT FILING DATE: 1998-10-01  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: FastSeq for Windows Version 3.0  
Q ID NO 26  
LENGTH: 1341  
TYPE: DNA  
ORGANISM: Vespula vulgaris  
9-166-205B-26

Query Match 3.1%; Score 43.4; DB 4; Length 1341;  
Best Local Similarity 50.9%; Pred. No. 0.0025;  
Matches 141; Conservative 0; Mismatches 121; Indels 15; Gaps 1;  
396 TGGCTGTAACAACTCTCCGTTGTTGGTCTGAGTGGCTTATTTATGATGTTCTCAT 455  
455 TGGCTGTAACAACTCTCCGTTGTTGGTCTGAGTGGCTTATTTATGATGTTCTCAT 455  
456 GAAAAAATTTGAATATTTCCCTTCTAAAGTGCACCTGATGGCCACAGCTTGGAGCACA 515  
515 AAAACACTATAAATCTCGATGGCAATATACGATTAAATGGACATAGCTTAGGAGCACA 574  
516 CCTGGCTGGGAAGCTGGGTCAAGGATACAGGCTTGGAG-----AAT 560  
575 TGGCTTCAAGTTTGGAGGCAAAAGGTTCAAGATTAATAGGAAAAATATTTCTGAAT 634  
561 AACTGGTTGGACCCAGCTGGCCATTTTCCACACACTCCAAAGAGTCAAGCTAGA 620  
635 TATTGGCTTGGTCTGCTAGGCTTGGTTCGATTCAATTCATTTCCGAAGACTCTG 694  
621 CCCCTCGATGCCAACTTTGTTGACGTTATTCATACA 657  
695 CGAGACAGATGCAGATATGTTCAATTTATACATACA 731

LT 13  
US94-02629-26  
Sequence 26, Application PC/TUS9402629  
GENERAL INFORMATION:  
APPLICANT: King, Te-Piao  
TITLE OF INVENTION: CLONING AND RECOMBINANT PRODUCTION OF  
TITLE OF INVENTION: VESPID VENOM ENZYMES, SUCH AS PHOSPHOLIPASE AND  
TITLE OF INVENTION: HYALURONIDASE, AND IMMUNOLOGICAL THERAPIES BASED THEREON  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/02629  
FILING DATE: 10-MAR-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/180,209  
FILING DATE: 11-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/031,400  
FILING DATE: 11-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-074 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1341 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 153..1052  
PCT-US94-02629-26

Query Match 3.1%; Score 43.4; DB 5; Length 1341;  
Best Local Similarity 50.9%; Pred. No. 0.0025;  
Matches 141; Conservative 0; Mismatches 121; Indels 15; Gaps 1;  
Qy 396 TGGCTGTAACAACTCTCCGTTGTTGGTCTGAGTGGCTTATTTATGATGTTCTCAT 455  
Db 455 TGGCTGTAACAACTCTCCGTTGTTGGTCTGAGTGGCTTATTTATGATGTTCTCAT 455  
Qy 456 GAAAAAATTTGAATATTTCCCTTCTAAAGTGCACCTGATGGCCACAGCTTGGAGCACA 515  
Db 515 AAAACACTATAAATCTCGATGGCAATATACGATTAAATGGACATAGCTTAGGAGCACA 574  
Qy 516 CCTGGCTGGGAAGCTGGGTCAAGGATACAGGCTTGGAG-----AAT 560  
Db 575 TGGCTTCAAGTTTGGAGGCAAAAGGTTCAAGATTAATAGGAAAAATATTTCTGAAT 634  
Qy 561 AACTGGTTGGACCCAGCTGGCCATTTTCCACACACTCCAAAGAGTCAAGCTAGA 620  
Db 635 TATTGGCTTGGTCTGCTAGGCTTGGTTCGATTCAATTCATTTCCGAAGACTCTG 694  
Qy 621 CCCCTCGATGCCAACTTTGTTGACGTTATTCATACA 657  
Db 695 CGAGACAGATGCAGATATGTTCAATTTATACATACA 731

RESULT 14  
US-08-446-855A-1/c  
Sequence 1, Application US/08446855A  
Patent No. 5845573  
GENERAL INFORMATION:  
APPLICANT: Stewart, Thomas S  
APPLICANT: Flores, Maria V  
APPLICANT: O'Sullivan, William J  
TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl  
TITLE OF INVENTION: phosphate synthetase II  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye PC  
STREET: 1100 No. 5849573th Glebe Road, 8th Floor  
CITY: Arlington  
STATE: Virginia



COUNTRY: USA  
 ZIP: 22201-4714  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.24  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/446,855A  
 FILING DATE: 06-Jul-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Mitchard, Leonard C  
 REGISTRATION NUMBER: 29,009  
 REFERENCE/DOCKET NUMBER: 47-80  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 703-816-4000  
 TELEFAX: 703-816-4100  
 FORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 8920 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: genomic  
 -446-855A-1

Query Match 2.8%; Score 39.2; DB 2; Length 8920;  
 Best Local Similarity 46.4%; Pred. No. 0.15;  
 Matches 128; Conservative 0; Mismatches 148; Indels 0; Gaps 0;  
 763 TTAATTACACCTTTACTGAAATTTAACTTCAATGCTTACAAAAGAAATGGCTTCCTTC 822  
 8858 TTGTATATACAGTTCCTCTTTTTCGTTGAATTTTAAATAATTCATAACATTATA 8799  
 823 TTGACTGTAAACATGCCCGAAGTTATCAATTTTATGCTGAAAGCATTTCTTAATCCTGAT 882  
 8798 AATTATTATATAATTAAGAATTTATACATTTTAAATGTTTATATATTTTATTTT 8739  
 883 GCATTTATGCTTATCCTTTGTAGATCCTACACATCTTTTAAAGCAGGAATTCCTTT 942  
 8738 ATAAATTAATAATTTATTAATAATTTATTAATTTTGTAAATTCAGAAATCCATT 8679  
 943 TGTCCAAAGAGGTTGCCCAACATGGTCAATTTTGTGTAGATGATTTCACTTCAAAAAT 1002  
 8678 TTATACATATGATGCAATAATAAATTTGTATATATACAAATTTTAAATAATAA 8619  
 1003 ATGAAGACTAATGGATCATTATTTTAAACACA 1038  
 8618 ATCAATATAAATTAATGATCATATTTTAAATAAAAAACA 8583

Search completed: January 6, 2004, 11:12:28  
 Job time : 102 secs  
 P 15  
 -150-741-1/c  
 sence 1, Application US/09150741  
 ent No. 6183996  
 BRAL INFORMATION:  
 PLICANT: Stewart et al.  
 FILE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate  
 ent No. 6183996  
 FILE OF INVENTION: Synthetase II  
 JE REFERENCE:  
 RENT APPLICATION NUMBER: US/09/150,741  
 RENT FILING DATE: 1998-09-10  
 LIER APPLICATION NUMBER: PL6380  
 LIER FILING DATE: 1992-12-16  
 LIER APPLICATION NUMBER: AU93/00617  
 LIER FILING DATE: 1993-12-02  
 LIER APPLICATION NUMBER: 08/446,855  
 LIER FILING DATE: 1995-07-06  
 BER OF SEQ ID NOS: 15  
 TWARE: PatentIn Ver. 2.0  
 ID NO 1

LENGTH: 8920  
 TYPE: DNA  
 ORGANISM: Plasmodium falciparum  
 US-09-150-741-1  
 Query Match 2.8%; Score 39.2; DB 3; Length 8920;  
 Best Local Similarity 46.4%; Pred. No. 0.15;  
 Matches 128; Conservative 0; Mismatches 148; Indels 0; Gaps 0;  
 763 TTAATTACACCTTTACTGAAATTTAACTTCAATGCTTACAAAAGAAATGGCTTCCTTC 822  
 8858 TTGTATATACAGTTCCTCTTTTTCGTTGAATTTTAAATAATTCATAACATTATA 8799  
 823 TTGACTGTAAACATGCCCGAAGTTATCAATTTTATGCTGAAAGCATTTCTTAATCCTGAT 882  
 8798 AATTATTATATAATTAAGAATTTATACATTTTAAATGTTTATATATTTTATTTT 8739  
 883 GCATTTATGCTTATCCTTTGTAGATCCTACACATCTTTTAAAGCAGGAATTCCTTT 942  
 8738 ATAAATTAATAATTTATTAATAATTTATTAATTTTGTAAATTCAGAAATCCATT 8679  
 943 TGTCCAAAGAGGTTGCCCAACATGGTCAATTTTGTGTAGATGATTTCACTTCAAAAAT 1002  
 8678 TTATACATATGATGCAATAATAAATTTGTATATATACAAATTTTAAATAATAA 8619  
 1003 ATGAAGACTAATGGATCATTATTTTAAACACA 1038  
 8618 ATCAATATAAATTAATGATCATATTTTAAATAAAAAACA 8583



```
181 IPGLGRITGLDPAGFFPHNTPKVRLDPSDANFVDVHTNAARILFELGVGTIDACGHL 240
241 FYPNGGKHPGCCDLITPLKFNFNAYKKEMASFFDCNHRASYQFYAESILNPDFAIAYP 300
241 FYPNGGKHPGCCDLITPLKFNFNAYKKEMASFFDCNHRASYQFYAESILNPDFAIAYP 300
301 CRSYTSFKAGNCFKSCGECPTMGHPADRFHFNKMTNGSHYFLNTGSLSPFARWRHKL 360
301 CRSYTSFKAGNCFKSCGECPTMGHPADRFHFNKMTNGSHYFLNTGSLSPFARWRHKL 360
361 VKLSGSEVTQGVFLRVGGAGIKTGEBALVSGKLEPGMTYTKLIDADVNGNITSVQFI 420
361 VKLSGSEVTQGVFLRVGGAGIKTGEBALVSGKLEPGMTYTKLIDADVNGNITSVQFI 420
421 KXHLFEDSQKLGAEWINTSGKYGKSTFCSDIMGNILQNLKPC 467
421 KXHLFEDSQKLGAEWINTSGKYGKSTFCSDIMGNILQNLKPC 467

J.T. 2
J-038-517-2
Sequence 2, Application US/10038517
Publication No. US20020115844A1
GENERAL INFORMATION:
APPLICANT: Yu, Xuanchuan
PPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: No. US20020115844A1 Human Lipase and Polynucleotides Encoding b
FILE REFERENCE: LEX-0293-USA
CURRENT APPLICATION NUMBER: US/10/038, 517
CURRENT FILING DATE: 2002-01-03
PRIOR APPLICATION NUMBER: US 60/259, 830
PRIOR FILING DATE: 2001-01-05
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
2 ID NO 2
LENGTH: 467
TYPE: PRT
ORGANISM: homo sapiens
J-038-517-2

ary Match 100.0%; Score 2527; DB 14; Length 467;
Best Local Similarity 100.0%; Pred. No. 9.2e-245;
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MLGIWIVAFVFFGTSRGKVCYERLGCDFKGLPWTFTSTELVGLPWSPEKINTRFLIYT 60
1 MLGIWIVAFVFFGTSRGKVCYERLGCDFKGLPWTFTSTELVGLPWSPEKINTRFLIYT 60
61 IHNPNAYQEISAVNSSTIQASYFGTDKITRINIAGNKTGDKQWQDMCNVLLQLEDINCIN 120
61 IHNPNAYQEISAVNSSTIQASYFGTDKITRINIAGNKTGDKQWQDMCNVLLQLEDINCIN 120
121 LDWINGSREYTHAVNNLRVVGAEVAFIDVLMKKFEYSPSKVHLIGHSLGAHLAGEAGSR 180
121 LDWINGSREYTHAVNNLRVVGAEVAFIDVLMKKFEYSPSKVHLIGHSLGAHLAGEAGSR 180
181 IPGLGRITGLDPAGFFPHNTPKVRLDPSDANFVDVHTNAARILFELGVGTIDACGHL 240
181 IPGLGRITGLDPAGFFPHNTPKVRLDPSDANFVDVHTNAARILFELGVGTIDACGHL 240
241 FYPNGGKHPGCCDLITPLKFNFNAYKKEMASFFDCNHRASYQFYAESILNPDFAIAYP 300
241 FYPNGGKHPGCCDLITPLKFNFNAYKKEMASFFDCNHRASYQFYAESILNPDFAIAYP 300
301 CRSYTSFKAGNCFKSCGECPTMGHPADRFHFNKMTNGSHYFLNTGSLSPFARWRHKL 360
301 CRSYTSFKAGNCFKSCGECPTMGHPADRFHFNKMTNGSHYFLNTGSLSPFARWRHKL 360
361 VKLSGSEVTQGVFLRVGGAGIKTGEBALVSGKLEPGMTYTKLIDADVNGNITSVQFI 420
361 VKLSGSEVTQGVFLRVGGAGIKTGEBALVSGKLEPGMTYTKLIDADVNGNITSVQFI 420
421 KXHLFEDSQKLGAEWINTSGKYGKSTFCSDIMGNILQNLKPC 467
421 KXHLFEDSQKLGAEWINTSGKYGKSTFCSDIMGNILQNLKPC 467
```

```
421 KXHLFEDSQKLGAEWINTSGKYGKSTFCSDIMGNILQNLKPC 467

RESULT 3
US-10-403-745-3
Sequence 3, Application US/10403745
Publication No. US20030165975A1
GENERAL INFORMATION:
APPLICANT: Khodadoust, Mehran
APPLICANT: Kapeller-Libermann, Rosana
TITLE OF INVENTION: No. US20030165975A1 Human Lipase Proteins, Nucleic Acides Encod:
FILE REFERENCE: 10147-14
CURRENT APPLICATION NUMBER: US/10/403,745
CURRENT FILING DATE: 2003-03-31
PRIOR APPLICATION NUMBER: US/09/411,132A
PRIOR FILING DATE: 2000-09-12
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 467
TYPE: PRT
ORGANISM: Homo sapiens
US-10-403-745-3

Query Match 99.8%; Score 2523; DB 12; Length 467;
Best Local Similarity 99.8%; Pred. No. 2.3e-244;
Matches 466; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGIWIVAFVFFGTSRGKVCYERLGCDFKGLPWTFTSTELVGLPWSPEKINTRFLIYT 60
Db 1 MLGIWIVAFVFFGTSRGKVCYERLGCDFKGLPWTFTSTELVGLPWSPEKINTRFLIYT 60
QY 61 IHNPNAYQEISAVNSSTIQASYFGTDKITRINIAGNKTGDKQWQDMCNVLLQLEDINCIN 120
Db 61 IHNPNAYQEISAVNSSTIQASYFGTDKITRINIAGNKTGDKQWQDMCNVLLQLEDINCIN 120
QY 121 LDWINGSREYTHAVNNLRVVGAEVAFIDVLMKKFEYSPSKVHLIGHSLGAHLAGEAGSR 180
Db 121 LDWINGSREYTHAVNNLRVVGAEVAFIDVLMKKFEYSPSKVHLIGHSLGAHLAGEAGSR 180
QY 181 IPGLGRITGLDPAGFFPHNTPKVRLDPSDANFVDVHTNAARILFELGVGTIDACGHL 240
Db 181 IPGLGRITGLDPAGFFPHNTPKVRLDPSDANFVDVHTNAARILFELGVGTIDACGHL 240
QY 241 FYPNGGKHPGCCDLITPLKFNFNAYKKEMASFFDCNHRASYQFYAESILNPDFAIAYP 300
Db 241 FYPNGGKHPGCCDLITPLKFNFNAYKKEMASFFDCNHRASYQFYAESILNPDFAIAYP 300
QY 301 CRSYTSFKAGNCFKSCGECPTMGHPADRFHFNKMTNGSHYFLNTGSLSPFARWRHKL 360
Db 301 CRSYTSFKAGNCFKSCGECPTMGHPADRFHFNKMTNGSHYFLNTGSLSPFARWRHKL 360
QY 361 VKLSGSEVTQGVFLRVGGAGIKTGEBALVSGKLEPGMTYTKLIDADVNGNITSVQFI 420
Db 361 VKLSGSEVTQGVFLRVGGAGIKTGEBALVSGKLEPGMTYTKLIDADVNGNITSVQFI 420
QY 421 KXHLFEDSQKLGAEWINTSGKYGKSTFCSDIMGNILQNLKPC 467
Db 421 KXHLFEDSQKLGAEWINTSGKYGKSTFCSDIMGNILQNLKPC 467

RESULT 4
US-10-312-088-23
Sequence 23, Application US/10312088
Publication No. US20030219862A1
GENERAL INFORMATION:
APPLICANT: Agarwal, Pankaj
APPLICANT: Cogswell, John P.
APPLICANT: Kabanic, Karen S.
APPLICANT: Lai, Ying-Ta
APPLICANT: Martensen, Shelby A.
```

PLICANT: Murdock, Paul R.  
 PLICANT: Smith, Randall F.  
 PLICANT: Strum, Jay C.  
 PLICANT: Xiang, Zhaoxing  
 PLICANT: Xie, Qing  
 PLICANT: Rizni, Safia K.  
 TITLE OF INVENTION: NOVEL COMPOUNDS  
 LE REFERENCE: GP50029  
 PRESENT APPLICATION NUMBER: US/10/312,088  
 PRESENT FILING DATE: 2000-12-20  
 PRESENT APPLICATION NUMBER: PCT/US01/19929  
 PRESENT FILING DATE: 2001-06-22  
 PRESENT APPLICATION NUMBER: 60/213,161  
 PRESENT FILING DATE: 2000-06-22  
 PRESENT APPLICATION NUMBER: 60/213,156  
 PRESENT FILING DATE: 2000-06-22  
 NUMBER OF SEQ ID NOS: 44  
 ID NO 23  
 FTWARE: FastSeq for Windows Version 4.0  
 LENGTH: 461  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 -312-088-23

Query Match 90.1%; Score 2276; DB 12; Length 461;  
 Best Local Similarity 91.5%; Pred. No. 1.4e-219;  
 Matches 431; Conservative 8; Mismatches 18; Indels 14; Gaps 3;  
 QY 1 MGIWIVAFPGTSGRKEVCYERLGCPCDGLPWTRTFSTELVGLPWSPEKINTRELLYT 60  
 DB 1 MGIWIVAFPGTSGRKEVCYERLGCPCDGLPWTRTFSTELVGLPWSPEKINTRELLYT 60  
 QY 61 IHNPNAYQBEISAVNSTIQASVFGTDKITRINIAAGTKDQK--WORDMCMVLLQLEDINCIN 120  
 DB 61 IHNPNAYQBEISAVNSTIQASVFGTDKITRINIAAGTKDQK--WORDMCMVLLQLEDINCIN 120  
 QY 121 LDWINGSREYIHAVNNLRVGAEVAFYDVLMMKFEPSPKSVHLIGHSLGAHLAGEAGSR 180  
 DB 121 LDWINGSREYIHAVNNLRVGAEVAFYDVLMMKFEPSPKSVHLIGHSLGAHLAGEAGSR 180  
 QY 181 IGLGRITGLDPAGPFPHTPKVRLDPSDANFVDVHTNAARILFELGVGTIDACGHL 240  
 DB 181 IGLGRITGLDPAGPFPHTPKVRLDPSDANFVDVHTNAARILFELGVGTIDACGHL 240  
 QY 241 FYNGGKMPGCCEDLITPLKFNFNAYKEMASFFDCNHARSYQFYAESILNPDATAYP 300  
 DB 241 FYNGGKMPGCCEDLITPLKFNFNAYKEMASFFDCNHARSYQFYAESILNPDATAYP 300  
 QY 301 CRSYTSFKAGNCFPCSGKEGCPMGHAFDRPHFKMKNKNGSHYFLNTG---SLSPFARWR 356  
 DB 301 CRSYTSFKAGNCFPCSGKEGCPMGHAFDRPHFKMKNKNGSHYFLNTG---SLSPFARWR 356  
 QY 357 HKLSVKLSGSEVTQGTVFLRVGGAIGKTGEFAIVSGKLEPGMTYTKLIDADVNVGNITSV 416  
 DB 351 HKLSVKLSGSEVTQGTVFLRVGGAIGKTGEFAIVSGKLEPGMTYTKLIDADVNVGNITSV 410  
 QY 417 QFIWKHLFEDSQNKLGAEMWINTSGKYGYKSTFCSDQIMGNILQNLKPC 467  
 DB 411 QFIWKHLFEDSQNKLGAEMWINTSGKYGYKSTFCSDQIMGNILQNLKPC 461

US-09-735-933-4  
 Sequence 4, Application US/09735933  
 Patent No. US20020052034A1  
 GENERAL INFORMATION:  
 APPLICANT: GUEGLER, Karl et al  
 TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC  
 ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES  
 THEREOF  
 FILE REFERENCE: CL000863  
 CURRENT APPLICATION NUMBER: US/09/735,933  
 CURRENT FILING DATE: 2000-12-14  
 NUMBER OF SEQ ID NOS: 6  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 5  
 LENGTH: 467  
 TYPE: PRT  
 ORGANISM: Mus musculus  
 US-09-735-933-5

Query Match 46.7%; Score 1180; DB 9; Length 467;  
 Best Local Similarity 49.0%; Pred. No. 1.7e-109;  
 Matches 231; Conservative 73; Mismatches 159; Indels 8; Gaps 6;  
 QY 1 MGIWIVAFPGTSGRKEVCYERLGCPCDGLPWTRTFSTELVGLPWSPEKINTRELLYT 60  
 DB 1 MLILWTIPLFLGAAQGCYCNLGCFCSDAEFPWAGTAIRPLKLLPWSPEKINTRELLYT 60

NUMBER OF SEQ ID NOS: 6  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 4  
 LENGTH: 469  
 TYPE: PRT  
 ORGANISM: Myocastor\_coypus  
 US-09-735-933-4  
 Query Match 47.1%; Score 1191; DB 9; Length 469;  
 Best Local Similarity 49.8%; Pred. No. 1.3e-110;  
 Matches 236; Conservative 63; Mismatches 163; Indels 12; Gaps 7;  
 QY 1 MGIWIVAFPGTSGRKEVCYERLGCPCDGLPWTRTFSTELVGLPWSPEKINTRELLYT 60  
 DB 1 MLFVMTTGLLLATARGNEVCYSHLGCFCDEKPMAGTLQRPVKSLLPASPEINTRFLLYT 60  
 QY 61 IHNPNAYQBEISAVNSTIQASVFGTDKITRINIAAGTKDQK--WORDMCMVLLQLEDINC 118  
 DB 61 NENPNYQILITATDPATIKASNFNLHRTFVINGFDNGEKDWTIDICKMPQVEKVC 120  
 QY 119 INLDWINGSRE-YIHAVNNLRVGAEVAFYDVLMMKFEPSPKSVHLIGHSLGAHLAGEA 177  
 DB 121 ICVDWQGGSLAIYSQAVQNIQVGAEVAYLVQVLSDLQGYKPGNVHMGHSLGAHTAAEA 180  
 QY 178 GSRIPGL-GRITGLDPAGPFPHTPKVRLDPSDANFVDVHTNAARILFELGVGTIDAC 236  
 DB 181 GRRKGLVGRITGLDPAGPFPHTPKVRLDPSDANFVDVHTNAARILFELGVGTIDAC 240  
 QY 237 GHLDPYFNGGKMPGCC-GLITPLKFNFNAYKEMASFFDCNHARSYQFYAESILNPD 295  
 DB 241 GMDPFPNGGKMPGCC-GLITPLKFNFNAYKEMASFFDCNHARSYQFYAESILNPD 298  
 QY 296 FIAPPCRSYTSFKAGNCFPCSGKEGCPMGHAFDRPHFKMKNKNG--SHYFLNTGSLSP 353  
 DB 299 FLGYPCASYEEFQXGCGPCPAGCPCPKGHYADQF---QKGANGVETVFLNTGSDSNFP 355  
 QY 354 RWRKLSVKLSGSEVTQGTVFLRVGGAIGKTGEFAIVSGKLEPGMTYTKLIDADVNVGN 413  
 DB 356 RWRKLSVKLSGSEVTQGTVFLRVGGAIGKTGEFAIVSGKLEPGMTYTKLIDADVNVGN 415  
 QY 414 TSVOFIWKHLFEDSQNKLGAEMWINTSGKYGYKSTFCSDQIMGNILQNLKPC 467  
 DB 416 QKVKFLMHNNGINLLQPLGASQITVQSGEYGTNYKFCSSNTVQEDVQLSLSPC 469

RESULT 6  
 US-09-735-933-5  
 Sequence 5, Application US/09735933  
 Patent No. US20020052034A1  
 GENERAL INFORMATION:  
 APPLICANT: GUEGLER, Karl et al  
 TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC  
 ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES  
 THEREOF  
 FILE REFERENCE: CL000863  
 CURRENT APPLICATION NUMBER: US/09/735,933  
 CURRENT FILING DATE: 2000-12-14  
 NUMBER OF SEQ ID NOS: 6  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 5  
 LENGTH: 467  
 TYPE: PRT  
 ORGANISM: Mus musculus  
 US-09-735-933-5

Query Match 46.7%; Score 1180; DB 9; Length 467;  
 Best Local Similarity 49.0%; Pred. No. 1.7e-109;  
 Matches 231; Conservative 73; Mismatches 159; Indels 8; Gaps 6;  
 QY 1 MGIWIVAFPGTSGRKEVCYERLGCPCDGLPWTRTFSTELVGLPWSPEKINTRELLYT 60  
 DB 1 MLILWTIPLFLGAAQGCYCNLGCFCSDAEFPWAGTAIRPLKLLPWSPEKINTRELLYT 60



```

T 9
-735-933-6
ence 6, Application US/09735933
ent No. US20020052034A1
ERAL INFORMATION:
PLICANT: GUEGLER, Karl et al
LE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
LE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
LE OF INVENTION: THEREOF
LE REFERENCE: CL000863
RENT APPLICATION NUMBER: US/09/735,933
RENT FILING DATE: 2000-12-14
MBER OF SEQ ID NOS: 6
WARE: FastSeq for Windows Version 4.0
ID NO 6
ENGTH: 467
YPE: PRT
RGANISM: Rattus_norvegicus
-735-933-6

ry Match 46.4%; Score 1172; DB 9; Length 467;
t Local Similarity 49.3%; Pred. No. 1.1e-108;
ches 232; Conservative 73; Mismatches 158; Indels 8; Gaps 6;

1 MLGIWIVAFLEFCTSRGKEVCYVERLGCFCFKDGLPWTFTSTELVGLPWSPEKINTREFLLYT 60
1 MLTLWTVSLFLGAAQGEVCYVDNLGCFSDAEPWAGTAIRPLKLLPWSPEKINTREFLLYT 60
61 IHNPNAYQEISAVNSSTIOASYFGTDKITRINIAGW--KTDGKWQORDMNCVLLQLEDINC 118
61 NENPTAFQTLQSDPLTIGASNFOVARKTRFIHGFIDKGEENWVVDCKNMFQVEVNC 120
119 INLDWINGSR-EYIHAVNNLRVGAEVAYFIDVLMKKFEYSPSKVHLIGHSLGAHLAGEA 177
121 ICVDWKKGSQTTYTOAANNVRVGAQVAMIDILVKNSYSPSKVHLIGHSLGAHVAGEA 180
178 GSRIPGLGRITGLDPAGPFPHNTPEKVRLEDPDANFVDVHTNAARILFELGVGTIDACG 237
181 GSRTPLGLGRITGLDPVEANFEGTPEEVRLEDPDADFVDVHTDAAPLIPFLGFTNOMSG 240
238 HLDYFNGGKHPGC-EDLITPLKFNFNAYKEMASFFDCNHARSYQFYAESILNPDFA 296
241 HLDFFNGGQMPGCKKVALSQIV--DIDGIWSGTRDFVACNHLRSYKYLESILNPDGF 298
297 IAYPCRSYTSFKAGNCFCKGCGPTMGHFADRPHFKMKTNGSHYFLNTGSLSPFARWR 356
299 AAYPCASYKDFESNCKFCPCDQCGPQMGHYADKFAGKS-GDEPOKFFLNTGEAKNFARWR 357
357 HKLSVKLSGSEVTQGTVFLRVGGAIGKTGEFAIVSGKLEPGMTYTKLIDADVNGNITSV 416
358 YRVSLILSGRMT-GQVKVALFGSKGNTRQYDIFRGIKPGATHSSSEFADKLDVGTIEKV 416
417 QFTWKXHLFEDSQNKLGAEWINTSGKYKYSTFCSDODIMGPNILQNKPC 467
417 KFLWNNQVINPSPFKVGAAKITVQKGEERTYFNFCSEETVREDTLLTL LPC 467

T 10
-403-745-8
ence 8, Application US/10403745
lication No. US20030165975A1
ERAL INFORMATION:
PLICANT: Khodadoust, Mehran
PLICANT: Kapeller-Libermann, Rosana
LE OF INVENTION: No. US20030165975A1el Human Lipase Proteins, Nucleic Acids Encod
LE OF INVENTION: Them, and Uses of Both of These
LE REFERENCE: 10147-14
RENT APPLICATION NUMBER: US/10/403,745
RENT FILING DATE: 2003-03-31
IOR FILING DATE: 2000-09-12
MBER OF SEQ ID NOS: 10

```

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-10-403-745-8

Query Match 46.4%; Score 1172; DB 12; Length 473;
Best Local Similarity 49.3%; Pred. No. 1.1e-108;
Matches 232; Conservative 73; Mismatches 158; Indels 8; Gaps 6;

QY 1 MLGIWIVAFLEFCTSRGKEVCYVERLGCFCFKDGLPWTFTSTELVGLPWSPEKINTREFLLYT 60
Db 1 MLTLWTVSLFLGAAQGEVCYVDNLGCFSDAEPWAGTAIRPLKLLPWSPEKINTREFLLYT 60
QY 61 IHNPNAYQEISAVNSSTIOASYFGTDKITRINIAGW--KTDGKWQORDMNCVLLQLEDINC 118
Db 61 NENPTAFQTLQSDPLTIGASNFOVARKTRFIHGFIDKGEENWVVDCKNMFQVEVNC 120
QY 119 INLDWINGSR-EYIHAVNNLRVGAEVAYFIDVLMKKFEYSPSKVHLIGHSLGAHLAGEA 177
Db 121 ICVDWKKGSQTTYTOAANNVRVGAQVAMIDILVKNSYSPSKVHLIGHSLGAHVAGEA 180
QY 178 GSRIPGLGRITGLDPAGPFPHNTPEKVRLEDPDANFVDVHTNAARILFELGVGTIDACG 237
Db 181 GSRTPLGLGRITGLDPVEANFEGTPEEVRLEDPDADFVDVHTDAAPLIPFLGFTNOMSG 240
QY 238 HLDYFNGGKHPGC-EDLITPLKFNFNAYKEMASFFDCNHARSYQFYAESILNPDFA 296
Db 241 HLDFFNGGQMPGCKKVALSQIV--DIDGIWSGTRDFVACNHLRSYKYLESILNPDGF 298
QY 297 IAYPCRSYTSFKAGNCFCKGCGPTMGHFADRPHFKMKTNGSHYFLNTGSLSPFARWR 356
Db 299 AAYPCASYKDFESNCKFCPCDQCGPQMGHYADKFAGKS-GDEPOKFFLNTGEAKNFARWR 357
QY 357 HKLSVKLSGSEVTQGTVFLRVGGAIGKTGEFAIVSGKLEPGMTYTKLIDADVNGNITSV 416
Db 358 YRVSLILSGRMT-GQVKVALFGSKGNTRQYDIFRGIKPGATHSSSEFADKLDVGTIEKV 416
QY 417 QFTWKXHLFEDSQNKLGAEWINTSGKYKYSTFCSDODIMGPNILQNKPC 467
Db 417 KFLWNNQVINPSPFKVGAAKITVQKGEERTYFNFCSEETVREDTLLTL LPC 467

RESULT 11
US-10-403-745-10
; Sequence 10, Application US/10403745
; Publication No. US20030165975A1
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: No. US20030165975A1el Human Lipase Proteins, Nucleic Acids Encodi
; TITLE OF INVENTION: Them, and Uses of Both of These
; FILE REFERENCE: 10147-14
; CURRENT APPLICATION NUMBER: US/10/403,745
; PRIOR FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US/09/411,132A
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Canis familiaris
; US-10-403-745-10

Query Match 46.1%; Score 1166; DB 12; Length 467;
Best Local Similarity 48.0%; Pred. No. 4.3e-108;
Matches 226; Conservative 79; Mismatches 158; Indels 8; Gaps 6;

QY 1 MLGIWIVAFLEFCTSRGKEVCYVERLGCFCFKDGLPWTFTSTELVGLPWSPEKINTREFLLYT 60
Db 1 MVSINTIALFLLGAQKAEVCYEQIGCFSDAEPWAGTAIRPLKVLFPWSPEKINTREFLLYT 60

```

```

61 IHNPNAYOEI SAVNSSTIOASVYFGTDKIIFINTAGW--KTDGKWQEDMCNVLLQLEDINC 118
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
61 NKQNPNNFTLDPDPSFTEIASNFQTKDKTFRFIILGHFDIGKEENWLDMCKNPKFVEEVC 120
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
119 INLDWINGSR-EYIHAVNNRVVGAEVAYIDVLMMKFEPSPSKHLIGHSLCAHLAGEA 177
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
121 ICVDWKKGQSQTSTAANNVRVVGAOVAQLMSLSANYSPSQVOLIGHSLGAHVAGEA 180
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
178 GSRIPLGLRGITGLDDPAGPPPHNTPKERVLRDPDSANFDVIHTNAARILFELGVGTIDAC 237
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
181 GSRTPGLGRITGLDDPVSEASFQGPTEEVRLPDTPDAFDVIHTDAABFLIPFLGFSGTSQQMG 240
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
238 HLDIFYPNGKHMPGC-EDLTPTLLKTFNFAYKKEMASFFDCNHARGYQFVASISLNPDAP 296
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
241 HLDFFPNNGGEEMPCKKNALSQIV--DLDGIMEGTRDFVACNHLRSVKYSISLNPDGF 298
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
297 IAYPCSYSTSFKAQNCFFCSKECQPTWGIFADRFHPKNNKTNGSHVFNLGTSLSPFAWR 356
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
299 ASYPCASRYAFESNKCFPCPDQCQPOMGHYADFAYKT-SDETQKYFLNLGTSSSNFAWR 357
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
357 HKLSVXLSGSEVTQGTVLFVRVGAIKGTGBEFAIVSGKLPGPMYTYKLIADVNVMGNTSV 416
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
358 YGVSIITLSGRKAT-GQAKVALFGSKGNTHQFNFKGILKPGSTHSNEFPDAKLDVGTIEKV 416
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
417 QFTWKXHLFEDSQNKLAGAEWINTSKYGVKSTFCSQDTMGNINLONLKPC 467
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
417 KFLWNNNVNVPTEPKYGAAKITVQKGEEXTVSCFSESIVREDBVLLTLTPC 467
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

```

LT 12  
0-403-745-4  
quence 4, Application US/10403745  
lication No. US20030165975A1  
NERAL INFORMATION:  
PPPLICANT: Khodadoust, Mehran  
PPPLICANT: Kapeller-Libermann, Rosana  
ITILE OF INVENTION: No. US20030165975A1el Human Lipase Proteins, Nucleic Acids Encod  
ITILE OF INVENTION: Them, and Uses of Both of These  
ITILE REFERENCE: 10147-14  
URRENT APPLICATION NUMBER: US/10/403,745  
URRENT FILING DATE: 2003-03-31  
RIOR APPLICATION NUMBER: US/09/411,132A  
RIOR FILING DATE: 2000-09-12  
UMBER OF SEQ ID NOS: 10  
FTWARE: PatentIn Ver. 2.1  
2 ID NO 4  
LENGTH: 467  
TYPE: PRT  
ORGANISM: Homo sapiens  
0-403-745-4

```

ary Match      46.0%; Score 1162; DB 12; Length 467;
st Local Similarity 48.6%; Pred. No. 1.1e-107;
ches 229; Conservative 75; Mismatches 159; Indels 8; Gaps 6;

1  MLGIWIVAFLLPFGTSRGCYCYVRLGCFVGLPWTRTSTELVGLPWSPEKINRFLPLVT 60
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1  MLIFWTITLFLGLGAAGKEVCYEDLGCFSDFTEPWGTAIRPLKILPWSPEKIGTRFLPLVT 60
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

61  IHNPNAYQETISAVNSSTIQASYPGTDKTIIRINAGW--KTDGKWQRDMQCNVLQLQEDINC 118
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
61  NENPNFNQIILLSDPSTIEASNFQMDRKTRFTIHGFDKGDSEWTDCKMLFEVEVNC 120
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

119  INLDWINGSR-EYIHAVNNLRVVVGAEVAVFIDVLMKKPEXSPSKVHLIGHSLCAHLAGEA 177
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
121  ICVDWKKGSQATYITQAANNRVVVGVAQVQMLDILLITEYSYPPSKVHLIGHSLGAHVAGEA 180
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

178  GSRIRGLGRITGLDPACPPHNTPKXYRLDPSDANFVDVHTNNAARILFELGVGTIDACG 237
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
181  GSKTPGUSRIITGLDPVEASPESTPEEVRLLDPSDAFDVDVHTDAAPLIFLFGFTNQMG 240
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

238  HLDVPYNGKGHPGOC-EDLITPLKFNFNAYKKNWASPFDCNARSQFYASSIILNPAF 296
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

```

Db      241 HDFFPNGSGSPGCKKQALSOIV-PLDGIWAGTRDFVACNHLRSYKYLESILNPDGF 298
Qy      297 IAYPCRSYTSFKAAGNCFPCSKEGCPTMGHPADRFHFQNMKTINGSHYFLNTGSLSPFARWR 356
Db      299 AAYPCTSYKSPESDKCFPCPDGCGPQMGHYADKFAGETSEEQ-QKFFLNTGEASNPARR 357
Qy      357 HKLSVKLSGSEVITQGTVELRVGGAGIKGTGFALVSGKLEPGMYTYTKLIDADVNVGNITSV 416
Db      358 YGVSYITLSGRVAT-GQIKVALFGNKGKNTHOYSIFRGILKPGSTHSYEFPAKLDVGTIEKY 416
Qy      417 QFIWKKHLPESQNKLGAEWINTSGKYGYKSTFCSDIMGPNIQLNKLKPC 467
Db      417 KELNNNVINPLPVGATKITVQKGEKTVYNFCSEDTVREDTLLTLTPC 467

RESULT 13
US-10-403-745-5
; Sequence 5, Application US/10403745
; Publication No. US20030165975A1
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; APPLICANT: Kapeller-Libermann, Roana
; TITLE OF INVENTION: No. US20030165975A1e1 Human Lipase Proteins, Nucleic Acids
; TITLE OF INVENTION: them, and Uses of Both of These
; FILE REFERENCE: 10147-14
; CURRENT APPLICATION NUMBER: US/10/403,745
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US/09/411,132A
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-403-745-5

```

Query Match	45.6%	Score 1152	DB 12	Length 469
Best Local Similarity	47.8%	Pred. No. 1.1e-106		
Matches 225	Conservative 68	Mismatches 172	Indels	Gaps 5
QY	1	MLGIWIAFLPFGSRGECVYCRLCGFKDGLPWTPTPTSELVGLPWPSEKINTREFLLYT	60	
Db	1	MLPPTWLGILLLATVRGKVCYQLGCFSDKEKPWAGTLQRPVKLLPWPSPEDIDTRFLYT	60	
QY	61	IHNPNWAOEISAVNSSTIQASITYGTDTKTRINIAGW--KTDGKWORDCMNVLLOEDLINC	118	
Db	61	NENPNNFQIITGTPDTEASNTQLODRKTRFIHGFELDKAEDSWPDMCKMFEVEKVC	120	
QY	119	INLQWJNSRE-YIHAVNNLRVVGAEVAFIDVLMKKFEPYSPKVVHLGSHGAHLGAEA	177	
Db	121	ICVDNRHGRSAMYQAVQNIIRVGAETAFLIQALSTQLGYSLEDVHVVLGSHGATAAEA	180	
QY	178	GSRIPG-LGRITGLDPAGPFPHNTKPEVRLDPSDFANFVDVIHTNAARILFELGVGTIDAC	236	
Db	181	GRLLGGRVGRITGLDPAGPCFQDEBEVRLDPSDAFVDVVIHTDSSPIVPSLFGMSQKV	240	
QY	237	GHLPDPYNGGKHEPGCEDLITPLLKNFNAYKEMASPDFCNHASYQFYAESILNPDAF	296	
Db	241	GHLPDPYNGGKHEPGCKKNVLSIT-DIGIIEGIGGVSCNHLRSFEYISSVLPDPGF	299	
QY	297	IAYPCRSYTFPKAGNCFFCSEKCEPTMGHFADRHFHPKNKMTGSHYFLNTLSLSEPAWR	356	
Db	300	LCYPCASYDEFOESKCPCCPAEGCPKNGHVADQPKGTSABEQT--FFLNTGESGNFTSMR	358	
QY	357	HKLIVKLSGSEVITGTVFLRVGGAIGKTEGPAIVSGKLBPGMTYTKLIDADVNVGNITSV	416	
Db	359	YKSVTVLSGKEKVNGYIRIAYIGSNENSKQIETFKGLSPDASHTKALDVDFNVGKIQKV	418	
QY	417	QFIWKGHLFDSQNLGAEMVINTSGKYKSTFCSDIMGPNIQLNKKPC	467	
Db	419	KFLNKKRGINLSEPKLCAISOITVOSGEDGTENVFCSSDTPVENVLIOSLYVP	469	



I 14  
-925-297-739  
ence 739, Application US/09925297  
ent No. US20020081659A1  
ERAL INFORMATION:  
PICANT: Rosen et al.  
LE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
RE REFERENCE: PA105  
RENT APPLICATION NUMBER: US/09/925,297  
RENT FILING DATE: 2001-08-10  
IOR APPLICATION NUMBER: PCT/US00/05989  
IOR FILING DATE: 2000-03-08  
IOR APPLICATION NUMBER: 60/124,270  
IOR FILING DATE: 1999-03-12  
MBER OF SEQ ID NOS: 928  
FTWARE: Patent in Ver. 2.0  
ID NO 739  
ENGTH: 477  
YPE: PRT  
RGANISM: Homo sapiens  
-925-297-739

ry Match 45.6%; Score 1152; DB 9; Length 477;  
t Local Similarity 47.8%; Pred. No. 1.1e-106;  
ches 225; Conservative 68; Mismatches 172; Indels 6; Gaps 5

1 MLCIGTAVLFFGTSRGKVCYERLGCDFGLPWTRIFSTELVCLPHSPKINTRLIYLT 60  
9 MLPFWLGLLLATVGRKEVCYGLGCFSDKRWAGTLQRPVKLLPSPEDIDTRLFLYT 68  
61 IHNPNAYQISAVNGSTIQASYFGTDKTRINIAGW--KTDGKQRDMCNVLQLQLEDINC 118  
69 NENPNNFQILITGPTDTEASNFQDRKTRFIHGFLLDKAEDSWPDMCKMFEVEKVC 128  
119 INLDWINGSR-EYIAVNNLNVVGAEVAFYDVLMMKFEPYSPSKVHLIGHSLGAHLAGEA 177  
129 ICVDWTHGRGAMYTQAVQIRVVGAETFLQALSTQLGYSLEDDVHVIGHSLGAHTAAEA 188  
178 GSRIPG-LGRITGLDPAGFPFHNTPEKVELRDPDSANFVDVIHTNAARILFELGVGTIDAC 236  
189 GRGLGRGVRITGLDPAGPCFQDEPEEVELRDPSDAVFVDVIHTDSSIPVPSLGFQMSQKV 248  
237 GHLDYFNGGKMGKBCEDLITPLLKFNENAYKKEMASFFQCNHARSQFVABSLINPDFAF 296  
249 GHLDFFNGKMPCKKNVLSTIT-DIGIWEIGGFVSCNHURSFYFYYSSSVLNPDDGF 307  
297 IAYPCRSYTSFXAGNCFCKSGEGCTMGHPADRFHFXNMKTNGSHVPLANTGSLSPFARWR 356  
308 LGYPCASYPDEFQESKCFPCPAGCGCPKMGHYADQFKGTSVAEQT--FFLNTGESGNFTSWR 366  
357 HKLSVKLSGSEVTTQGVFLVRVGAIGTGEFAIVSGKLEPGMYTKLIDADVNVGNITSV 416  
367 YKVSVTLGSKKGVYIRIALYGSNENSKQVEIFKGLKPDASHCAIDVDVFNGLQKV 426  
417 QFIWKHKLHEDSONKLGAEMVINTSGKYGKSTFCSDQIMGNILQNLKPC 467  
427 KFLNNKRGINTLSEPKLGASQITVQSGEDGTENFCSSDPTVEENVLOSILYPC 477

```

; CURRENT FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-923-779-151

Query Match
Best Local Similarity 44.1%; Score 1115; DB 9; Length 465;
Matches 226; Conservative 73; Mismatches 160; Indels 14; Gaps 10;

Qy 1 MLGIWIVAFLEFCTSRGKEYCYERLCQFKDGLPWTRTFTSELVGLPWSPEKINTRFLLYT 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MLPLWTLS-LLLCAGAVGKVCYERLCGCFSDSPSGITERPLHILPWSKDVNTRFLLYT 59

Qy 61 IHNPNAYOEISAVNSSTIQASVFGTDKTIIRINIAGH--KTDGKWQDMCNVLLQLEDINC 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 NENPNNFQEVAA-DSSSISGSNFKTRKRFIIHGFIIDGKEENWLANVCNLFKVESVNC 118

Qy 119 INLDWINGGRE-YIHAVNNLRVVGAEVAYFIDVLMKKFKEYSPSKVHLIGHSLGAHLAGEA 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 ICVDWKGSGRTGYTQASQIRIVGAEVAYFVEFLQSAFGYSPSNVHVIGHSLGAHAAGEA 178

Qy 178 GSRIPG-LGRITGLDPAGFFHTPKVELDPSDANFVDVITHNAARIILFELGVGTIDAC 236
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 GRNTGTIGRITGLDPAEFCFGTPELVLDPSDAKFVDVITHDGAPIVFNLFGMSQV 238

Qy 237 GHLDYFNGGKHPGCCEDLITPLLKFNFNAYKEMASFFDCNHARSYQPYAESILNPDAF 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 GHLDFFNGGVEMPGCKKMLSQI-VDIDGIWEGTRDFAACNHLRSYKYTYDSIVNPDGF 297

Qy 297 IAYPCRSYTSERAGNCFCKSGCPTMGHFADRFHFFKNKTN--GSHYFLNTGSLSPFAR 354
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 AGFPCCASYNVFTANKFCPCPGSGCPOMGHYADRY---PGKTDNVGQKFDYLDTGDAENFAR 354

Qy 355 WRHKLSQLSGSEVTOGTGYFLRVGGIAGTGBFAIVSGKLEPQMTYTKLIDADVNVGNIT 414
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 355 WRVKSVTLTSGKKVT--GHILVSLFGNKGNSKQYEIPKGTLPKDSYTHSNEFSDSDVDGDLQ 413

Qy 415 SVQPIWKHLFEDSQNLGAEMVINTSGYGVKSTFCODIMGPNILQNLKPC 467
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 414 MKVFIYNNVINPTLPRVGASKII-VETNVGQKFPCCSPETVREEVLLTLTPC 465

Search completed: January 6, 2004, 11:28:12
Job time : 53 secs

```

P 15  
-923-779-151  
ence 151, Application US/09923779  
nt No. US20020076721A1  
REAL INFORMATION:  
PLICANT: Pyle, Ruth A.  
PLICANT: Xu Jiangchun  
PLICANT: Kalos, Michael D.  
E OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
E OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER  
E REFERENCE: 210121.553  
RENT APPLICATION NUMBER: US/09/923,779

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

cleic - nucleic search, using sw model

n: January 6, 2004, 08:31:36 ; Search time 411 Seconds  
(without alignments)  
9221.438 Million cell updates/sec

ct score: 1404

nce: 1 agcctgaattgattgt.....agaacctgaaccatgtctaa 1404

ng table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

hed: 2552756 seqs, 1349719017 residues

number of hits satisfying chosen parameters: 5105512

um DB seq length: 0

um DB seq length: 2000000000

processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

ase : N\_Geneseq\_19Jun03:\*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

t	Score	Query Match	Length	DB	ID	Description
1	1404	100.0	1404	24	ABL59436	Nucleotide sequenc
2	1404	100.0	1422	24	ASN87363	Human lipase encod
3	1402.4	99.9	2352	22	RAD03848	Human lipase prote
4	1234.6	87.9	1383	24	AAD27797	Human pancreatic 1
5	1115.8	79.5	1526	24	ABL60540	Human lipid metabo
6	978.8	69.7	1389	24	ABA90354	Human polynucleoti
7	808	57.5	1540	24	ABN60002	Novel human coding
8	386.2	27.5	1481	24	ABL69985	Pancreas cancer re

9	353.8	25.2	1427	23	AAS81082	DNA encoding novel
10	342.6	24.4	1454	24	ABL59552	Human triacylglyce
11	342.6	24.4	1471	24	ABK44205	cDNA #145 encoding
12	342.6	24.4	1471	24	ABL69968	Pancreas cancer re
13	342.6	24.4	1536	21	ABL98880	Human pancreatic c
14	338	24.1	1410	23	AAS81083	DNA encoding novel
15	338	24.1	1450	24	ABL69986	Pancreas cancer re
16	338	24.1	1506	21	AAC99052	Human pancreatic c
17	334.8	23.8	1422	14	AAQ35056	GPL(+) mutant with
18	330.2	23.5	1404	14	AAQ35059	Human pancreatic 1
19	297.8	21.2	1206	21	AAC00081	Human secreted pro
20	276.4	19.7	1384	14	AAQ35058	Guinea pig pancrea
21	245	17.5	915	23	AAS83973	DNA encoding novel
22	245	17.5	915	23	AAS85111	DNA encoding novel
23	245	17.5	915	23	AAS91515	DNA encoding novel
24	245	17.5	1142	23	AAS89469	DNA encoding novel
25	234	16.7	878	22	AAI89392	Human polynucleoti
26	224	16.0	938	24	ABK44199	CDNA #139 encoding
27	209.2	14.9	907	24	ABK44136	CDNA #76 encoding
28	204.2	14.5	879	24	ABK44154	CDNA #94 encoding
29	203	14.5	963	24	ABK44082	CDNA #22 encoding
30	202.6	14.4	766	24	ABK44133	CDNA #73 encoding
31	198.8	14.2	872	24	ABK44164	CDNA #104 encoding
32	194.4	13.8	800	24	ABV95086	Human pancreatic c
33	192.6	13.7	789	24	ABV95051	Human pancreatic c
34	189.2	13.5	735	24	ABK44142	CDNA #82 encoding
35	188.6	13.4	633	21	AAC89794	Human pancreatic c
36	187	13.3	750	24	ABK44137	CDNA #77 encoding
37	186.2	13.3	764	24	ABK44138	CDNA #78 encoding
38	179.8	12.8	548	24	ABV97752	Human pancreatic c
39	179.8	12.8	567	24	ABV97965	Human pancreatic c
40	179.6	12.8	789	24	ABK44135	CDNA #75 encoding
41	179.4	12.8	599	24	ABV98006	Human pancreatic c
42	178.8	12.7	526	24	ABV97817	Human pancreatic c
43	178.8	12.7	780	24	ABK44145	CDNA #85 encoding
44	177.6	12.6	741	24	ABK44140	CDNA #80 encoding
45	175	12.5	761	24	ABK44134	CDNA #74 encoding

## ALIGNMENTS

RESULT 1  
ABL59436  
ID ABL59436 standard; cDNA; 1404 BP.

XX ABL59436;

XX 22-OCT-2002 (first entry)

XX Nucleotide sequence of a human lipase.

XX Human; lipase; enzyme; mental disorder; disease; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1404

FT /tag= a

FT /product= "lipase"

FT replace (378, G)

FT /tag= b

FT replace (1141, G)

FT /tag= c

FT /note= "this results in the mutation Ile381Val"

FT replace (1144, A)

FT /tag= d

FT /note= "this results in the mutation Gly382Arg"

PN WO200253753-A2.

XX 11-JUL-2002.

XX



Human; lipase; enzyme; therapeutic; gene; ss.

Homo sapiens.

Key Location/Qualifiers  
 CDS 9..1412  
 /\*tag= a  
 /product= "lipase"

US2002052034-A1.

02-MAY-2002.

14-DEC-2000; 2000US-0735933

28-SEP-2000; 2000US-235925P

(GUEG/) GUEGLER K.

(WEBS/) WEBSTER M.

(KETC/) KETCHUM K A.

(DFRA/) DI FRANCESCO V.

(BEAS/) BEASLEY E M.

Guegler K, Webster M, Ketchum KA, Di Francesco V, Beasley EM;

WPI; 2002-443697/47.

P-PSDB; ABB79039.

A lipase protein identified from the human genome is related to the pancreatic lipase subfamily, and is useful to provide new drug targets in the discovery of therapeutics for lipase-mediated disease

Claim 4; Fig 1; 85pp; English.

The present sequence encodes a human lipase protein. Molecules from the present invention can be used as models for the development of human therapeutic targets, aid in the identification of therapeutic proteins and serve as targets for the development of human therapeutic agents that modulate lipase activity in cells and tissues that express lipase. N.B. The present sequence has been given as a protein sequence in the Sequence Listing of the present invention, but if you put the protein sequence into single letter code you get the nucleotide sequence given in Fig 1.

Sequence 1422 BP; 418 A; 279 C; 306 G; 419 T; 0 other;

ry Match 100.0%; Score 1404; DB 24; Length 1422;

t Local Similarity 100.0%; Pred.No. 0;

ches 1404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCTGGAAATTGGATTGCTTGCATCTTGTCTTTGGCAGATCAAGAGGAAAGATT 60

9 ATGCTGGAAATTGGATTGCTTGCATCTTGTCTTTGGCAGATCAAGAGGAAAGATT 68

61 TGCATGAAGGTTAGGTTGTTTCAAGATGTTTACCATGGACGAGACTTCTCAACA 120

65 TGCATGAAGGTTAGGTTGTTTCAAGATGTTTACCATGGACGAGACTTCTCAACA 128

121 GAGTTGGTAGTTTACCTGGTCTCAGAGAAGATAAACACTCGTTTCTGCTCACT 180

129 GAGTTGGTAGTTTACCTGGTCTCAGAGAAGATAAACACTCGTTTCTGCTCACT 188

181 ATACAAATCCCAATGCCCTATCAGGAGATCAGTGGGTTAATCTTCAACTATCCAGGC 240

189 ATACAAATCCCAATGCCCTATCAGGAGATCAGTGGGTTAATCTTCAACTATCCAGGC 248

241 TCATATTTGGAAACAGACAAGATCACCGGTATCAACATAGTGAATGGAACAGATGGC 300

249 TCATATTTGGAAACAGACAAGATCACCGGTATCAACATAGTGAATGGAACAGATGGC 308

301 AAATGSCAGAGACATGTGCAATGTGCTTACAGCTGGAGATATATATTCATTAAAT 360

Db 309 AATGCGCAGAGACATGTGCAATGTGTTGCTACAGCTGGAGATATAAATTCATTAAAT 368  
 QY 361 TTAGATTGATCAACGGTTACGGGAATACATCCATGCTGTAAACAATCTCCGTGTGTT 420  
 Db 369 TTAGATTGATCAACGGTTACGGGAATACATCCATGCTGTAAACAATCTCCGTGTGTT 428  
 QY 421 GGTGCTGAGGTGGCTTATTTTATGATGTTCTCATGAATAAATTTGAATATTTCCCTTCT 480  
 Db 429 GGTGCTGAGGTGGCTTATTTTATGATGTTCTCATGAATAAATTTGAATATTTCCCTTCT 488  
 QY 481 AAAGTGCACCTGATTGGCCACAGCTTGGAGCACACCTGGCTGGGGAGCTGGGTCAAGG 540  
 Db 489 AAAGTGCACCTGATTGGCCACAGCTTGGAGCACACCTGGCTGGGGAGCTGGGTCAAGG 548  
 QY 541 ATACCAGGCTTGGAGGAATTAACCTGGGTTGACCCAGCTGGGCCATTTTCCACAACACT 600  
 Db 549 ATACCAGGCTTGGAGGAATTAACCTGGGTTGACCCAGCTGGGCCATTTTCCACAACACT 608  
 QY 601 CCAAGGAAGTCAGCTAGACCCCTCGGATGCCCAACTTTGTTGACGTTATTCATACAAAT 660  
 Db 609 CCAAGGAAGTCAGCTAGACCCCTCGGATGCCCAACTTTGTTGACGTTATTCATACAAAT 668  
 QY 661 GCAGCTCGATCCTCTTTGAGCTTGGTGTGGAAACCATGATGCTTGTGTCATCTTGAC 720  
 Db 669 GCAGCTCGATCCTCTTTGAGCTTGGTGTGGAAACCATGATGCTTGTGTCATCTTGAC 728  
 QY 721 TTTTACCCAAATGGAGGAAGCACATGCCAGATGTGAAGACTTAATTACACCTTTACTG 780  
 Db 729 TTTTACCCAAATGGAGGAAGCACATGCCAGATGTGAAGACTTAATTACACCTTTACTG 788  
 QY 781 AAATTAACCTCAATGCTTACAAAAGAAATGGCTTCTTTGAGCTGTGAACCATGCC 840  
 Db 789 AAATTAACCTCAATGCTTACAAAAGAAATGGCTTCTTTGAGCTGTGAACCATGCC 848  
 QY 841 CGAAGTATCAATTTTATGCTGAAGCAATCTTAATCTGATGCATTTATTTGCTTATCCT 900  
 Db 849 CGAAGTATCAATTTTATGCTGAAGCAATCTTAATCTGATGCATTTATTTGCTTATCCT 908  
 QY 901 TGTAGATCCTACACATCTTTTAAAGCAGGAATTTGCTTCTTTGTTCCAAAGAGGTGC 960  
 Db 909 TGTAGATCCTACACATCTTTTAAAGCAGGAATTTGCTTCTTTGTTCCAAAGAGGTGC 968  
 QY 961 CCAACAATGGCTCAATTTGCTGATAGATTTCACTTCAAAAATATGAAGACTAATGGATCA 1020  
 Db 969 CCAACAATGGCTCAATTTGCTGATAGATTTCACTTCAAAAATATGAAGACTAATGGATCA 1028  
 QY 1021 CATTATTTTAAACACAGAGGTCCCTTTTCCCATTTTCCCGTTGGAGGCACAAATGTCT 1080  
 Db 1029 CATTATTTTAAACACAGAGGTCCCTTTTCCCATTTTCCCGTTGGAGGCACAAATGTCT 1088  
 QY 1081 GTTAACCTCAGTGAAGGAAGTCACTCAAGGAACCTGCTTCTTCTGTTAGCGGGGCA 1140  
 Db 1089 GTTAACCTCAGTGAAGGAAGTCACTCAAGGAACCTGCTTCTTCTGTTAGCGGGGCA 1148  
 QY 1141 ATTGGAAAACTGGGAGTTTGCATTTGTCAGTGGAAAACTTGAGCCAGGCATGACTTAC 1200  
 Db 1149 ATTGGAAAACTGGGAGTTTGCATTTGTCAGTGGAAAACTTGAGCCAGGCATGACTTAC 1208  
 QY 1201 ACAAAATTAATCGATGAGATGTTAAAGTTGGAAAAACATTAACAAGTGTTCAGTTCACTGG 1260  
 Db 1209 ACAAAATTAATCGATGAGATGTTAAAGTTGGAAAAACATTAACAAGTGTTCAGTTCACTGG 1268  
 QY 1261 AAAAAACATTTGTTTGAAGATTTCTCAGAAATAAGTTGGAGCAGAAATGGTGATATAACA 1320  
 Db 1269 AAAAAACATTTGTTTGAAGATTTCTCAGAAATAAGTTGGAGCAGAAATGGTGATATAACA 1328  
 QY 1321 TCTGGAAATATGATATAAATCTACCTTCTGTAGCAAGACATTTATGGACCTAATATT 1380  
 Db 1329 TCTGGAAATATGATATAAATCTACCTTCTGTAGCAAGACATTTATGGACCTAATATT 1388  
 QY 1381 CTCAGACCTGAAACCATGCTAA 1404  
 Db 1389 CTCAGACCTGAAACCATGCTAA 1412

LT 3

3848

AAD03848 standard; cDNA; 2352 BP.

AAD03848;

19-JUN-2001 (first entry)

Human lipase protein, MLip-1 cDNA.

Human; lipase; MLip-1; antiinflammatory; antilipemic; cardiant; therapy; anorectic; cholesterol biosynthesis; lipid metabolism; diabetes; obesity; pancreatic disorder; nutritional disorder; malabsorption; malnutrition; metabolic disorder; hyperlipidaemia; hypolipidaemia; abdominal lesion; lipidoses; pancreatic duct obstruction; Gaucher's disease; peritonitis; atherosclerosis; Niemann-Pick disease; arteriosclerosis; pancreatitis; coronary artery disease; perforated peptic ulcer; intestinal obstruction; linoleic acid deficiency; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 125..1528

/\*tag= a

/product= "Human lipase protein, MLip-1"

/note= "The coding region is specifically claimed

as SEQ ID NO:2 in claim 2 (page no: 80) of the

specification"

125..175

/\*tag= b

176..1525

/\*tag= c

/product= "Human mature lipase protein, MLip-1"

WO200125409-A1.

12-APR-2001.

02-OCT-2000; 2000WO-US27116.

01-OCT-1999; 99US-0411132.

(MILL-) MILLENIUM PHARM INC.

Khodadoust M, Kapeller-libermann R;

WPI: 2001-266302/27.

P-PSDB; AAE00508.

Novel isolated lipase, MLip-1, useful for identifying compounds which modulate MLip-1 activity, for treating conditions or disorders associated with aberrant activity of MLip-1 such as hypolipidemia, obesity

Claim 2; Page 95-96; 112pp; English.

The present sequence is human lipase protein, MLip-1 cDNA. MLip-1 catalyse the formation and cleavage of ester bonds between fatty acyl moieties and glyceride moieties. It is involved in dietary fat degradation and absorption, cholesterol biosynthesis and maintenance of plasma lipid and lipoprotein levels. MLip-1 is also useful for identifying compounds which modulates its activity and expression. These compounds are useful for preventing, diagnosing or treating disorders relating to inappropriate lipid metabolism and aberrant pancreatic function such as diabetes, obesity, nutritional disorders (e.g. lipid malabsorption and malnutrition), metabolic disorders (particularly including lipid metabolism anomalies such as hyperlipidaemia of types I to V and hypolipidaemia), pancreatitis, obstruction of the pancreatic duct, various lipidoses (e.g. Gaucher's disease and Niemann-Pick disease), atherosclerosis, arteriosclerosis, coronary artery disease, perforated peptic ulcer, abdominal lesions, intestinal obstruction,

CC peritonitis and linoleic acid deficiency. MLip-1 and its modulators can be used in screening assays, detection assays (e.g. chromosomal mapping, tissue typing and forensic biology assays), predictive medicine (e.g. CC diagnostic assays, prognostic assays, monitoring of clinical trials and pharmacogenomic applications) and methods of treatment (e.g. therapeutic CC and prophylactic methods).

XX Sequence 2352 BP; 705 A; 431 C; 491 G; 722 T; 3 other;

Query Match 99.9%; Score 1402.4; DB 22; Length 2352;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1403; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTTGGAAATTTGGATGTGTCATTCTTTTGGCACATCAAGAGGAAAAAGATT 60  
 DB 125 ATGCTTGGAAATTTGGATGTGTCATTCTTTTGGCACATCAAGAGGAAAAAGATT 184

QY 61 TGCATGAAAGGTTAGGGTGTTCCTCAAGATGGTTTACCATGACACAGGACTTCTCAACA 120  
 DB 185 TGCATGAAAGGTTAGGGTGTTCCTCAAGATGGTTTACCATGACACAGGACTTCTCAACA 244

QY 121 GAGTTGGTAGGTTTACCCCTGGTCTCCAGAGAGATAAACAACATCGTTTCCCTGCTACACT 180  
 DB 245 GAGTTGGTAGGTTTACCCCTGGTCTCCAGAGAGATAAACAACATCGTTTCCCTGCTACACT 304

QY 181 ATACACATCCCAATGCTATCAGGAGATCAGTGGGTTAATCTTCACTATCCAGCC 240  
 DB 305 ATACACATCCCAATGCTATCAGGAGATCAGTGGGTTAATCTTCACTATCCAGCC 364

QY 241 TCATATTTTGGAAACAGACAGATCACCCGTATCAACATAGCTGGATGAAAAACAGATGSC 300  
 DB 365 TCATATTTTGGAAACAGACAGATCACCCGTATCAACATAGCTGGATGAAAAACAGATGSC 424

QY 301 AAATGGCAGAGACATGTGCATGTGTCTACAGCTGGAGAGATATAAATTGCATTAAT 360  
 DB 425 AAATGGCAGAGACATGTGCATGTGTCTACAGCTGGAGAGATATAAATTGCATTAAT 484

QY 361 TTAGATTGGATCAACGGTTCAACGGGAATACATCCATGCTGTAAACAATCTCGTGTGTT 420  
 DB 485 TTAGATTGGATCAACGGTTCAACGGGAATACATCCATGCTGTAAACAATCTCGTGTGTT 544

QY 421 GGTGCTGAGTGCTTATTTTATGTGTTCTCATGAAAAAATTTGAATATTTCCCTTCT 480  
 DB 545 GGTGCTGAGTGCTTATTTTATGTGTTCTCATGAAAAAATTTGAATATTTCCCTTCT 604

QY 481 AAATGTCACCTTGATTCGCCACACAGCTTGGGAGACACCTGGCTGGGGAGAGCTGGGTCAAGG 540  
 DB 605 AAATGTCACCTTGATTCGCCACACAGCTTGGGAGACACCTGGCTGGGGAGAGCTGGGTCAAGG 664

QY 541 ATACCAGGCTTGGAAAGATAAATCTGGGTTGGACCCAGCTGGGCCATTTTCCACAACACT 600  
 DB 665 ATACCAGGCTTGGAAAGATAAATCTGGGTTGGACCCAGCTGGGCCATTTTCCACAACACT 724

QY 601 CCAAGGAGTCAAGCTAGACCCCTCGGATGCGCACTTGTGAGCTTATTCATACAAT 660  
 DB 725 CCAAGGAGTCAAGCTAGACCCCTCGGATGCGCACTTGTGAGCTTATTCATACAAT 784

QY 661 GCAGCTCGCATCCTCTTTCAGCTTGTGTGGAAACCATTTGATGCTTGTGGTCACTTTCAC 720  
 DB 785 GCAGCTCGCATCCTCTTTCAGCTTGTGTGGAAACCATTTGATGCTTGTGGTCACTTTCAC 844

QY 721 TTTTACCCAAATGGAGGAGACATGCCAGATGTGAAGACTTAATTACACCTTTACTG 780  
 DB 845 TTTTACCCAAATGGAGGAGACATGCCAGATGTGAAGACTTAATTACACCTTTACTG 904

QY 781 AAATTTAACTTCAATCTTACAAAAAAGAAATGGCTTCTCTTTTGAAGTAAACCATGCC 840  
 DB 905 AAATTTAACTTCAATCTTACAAAAAAGAAATGGCTTCTCTTTTGAAGTAAACCATGCC 964

QY 841 CGAAGTTATCAATTTTATCTGAAAGCATCTTAATCTGTATGCAATTTATGCTTATCCT 900  
 DB 965 CGAAGTTATCAATTTTATCTGAAAGCATCTTAAATCTGTATGCAATTTATGCTTATCCT 1024

901 TGAGATCCTACATCTTTAAAGCAGGAATTTGCTTTCTTTTCCAAAGAGTTGC 960  
 1025 TGAGATCCTACATCTTTAAAGCAGGAATTTGCTTTCTTTTCCAAAGAGTTGC 1084  
 961 CCAACAATGGTCAATTTGCTGATAGATTTCACTTCAAAAATATGAAGACTAATGGATCA 1020  
 1085 CCAACAATGGTCAATTTGCTGATAGATTTCACTTCAAAAATATGAAGACTAATGGATCA 1144  
 1021 CATTATTTTAAACACAGGTCCTTTTCCCATTTTCCCGTTGGAGGACAAATTTGCT 1080  
 1145 CATTATTTTAAACACAGGTCCTTTTCCCATTTTCCCGTTGGAGGACAAATTTGCT 1204  
 1081 GTTAACTAGTGAAGCGAAGTCACTCAAGGAACCTGCTTTCTTCTGTTAGGCGGGCA 1140  
 1205 GTTAACTAGTGAAGCGAAGTCACTCAAGGAACCTGCTTTCTTCTGTTAGGCGGGCA 1264  
 1141 ATTGGGAACACTGGGAGTTTGCATTTGCTGAGGAACTTCAAGGAACCTGCTTTCTTCTGTTAGGCGGGCA 1200  
 1265 ATTGGGAACACTGGGAGTTTGCATTTGCTGAGGAACTTCAAGGAACCTGCTTTCTTCTGTTAGGCGGGCA 1324  
 1201 AAAAAATTAATCGATCGAATGTTAAACGTTTGAACCAATTAACAAGTGTTCAGTTCACTCTGG 1260  
 1325 AAAAAATTAATCGATCGAATGTTAAACGTTTGAACCAATTAACAAGTGTTCAGTTCACTCTGG 1384  
 1261 AAAAAATTTGTTGAAGATTTCAAGATTAATAGTTGGGAGCAAGAAATGTTGAATAATACA 1320  
 1385 AAAAAATTTGTTGAAGATTTCAAGATTAATAGTTGGGAGCAAGAAATGTTGAATAATACA 1444  
 1321 TCTGGAAATATGATATAAATCTACTCTCTGTTAGCAAGACATTTAGGACCTTAATTT 1380  
 1445 TCTGGAAATATGATATAAATCTACTCTCTGTTAGCAAGACATTTAGGACCTTAATTT 1504  
 1381 CTCGAGAACTGAAACCATGCTAA 1404  
 1505 CTCGAGAACTGAAACCATGCTAA 1528

T 4

797

AAD27797 standard; DNA; 1383 BP.

AAD27797;

18-APR-2002 (first entry)

Human pancreatic lipase gene, sbg37163LIPASE.

Human; therapy; wound healing disorder; vaccine; cancer; infection; autoimmune disorder; haematopoietic disorder; inflammation; arthritis; Parkinson's disease; Huntington's chorea; schizophrenia; antiarrhythmic; multiple sclerosis; Alzheimer's disease; analgesic; cardiant; asthma; ischaemia; stroke; AIDS; bone disease; atherosclerosis; brain disorder; depression; cardiovascular disease; myocardial infarction; renal failure; respiratory disease; liver disorder; Fanconi's syndrome; spleen disorder; type II diabetes mellitus; skeletal muscle disorder; immunosuppressive; hyperplenism; renal disease; hypoglycaemia; gastrointestinal disease; neoptropic; cirrhosis; Hodgkin's disease; neuroleptic; antiinflammatory; haemostatic; vulnery; anticonvulsant; antirheumatic; neuroprotective; nephrotropic; hypotensive; vasotropic; cytostatic; cerebroprotective; allergy; pancreatic lipase; gene; ds.

Homo sapiens.

Key Location/Qualifiers

CDS

1..1383

/tag= a

/product= "Human pancreatic lipase"

/note= "CDS does not include stop codon"

/partial

W0200198342-A1.

27-DEC-2001.

XX 22-JUN-2001; 2001WO-US19929.  
 XX 22-JUN-2000; 2000US-213156P.  
 XX 22-JUN-2000; 2000US-213161P.  
 XX (SMIK ) SMITHKLINE BEECHAM CORP.  
 XX (SMIK ) SMITHKLINE BEECHAM PLC.  
 XX (GLAX ) GLAXO GROUP LTD.  
 XX Agarwal P, Cogswell JP, Kabnic KS, Lai Y, Martensen SA;  
 XX Murdock PR, Smith RF, Strum JC, Xiang Z, Xie Q, Rizni SK;  
 XX WPI; 2002-139783/18.  
 XX P-PSDB; AAE17302.  
 XX Novel secreted and membrane-associated polypeptides and polynucleotides useful for preventing, ameliorating or correcting dysfunction or disease including diabetes, cancer, hypertension and growth abnormalities  
 XX Claim 2; Page 83; 138pp; English.  
 XX The invention relates to secreted and membrane-associated polypeptides and polynucleotides. The sequences of the invention are useful in diagnostic assays for detecting diseases associated with inappropriate activity or levels of these polynucleotides, and in identifying their agonists and antagonists that are potentially useful in therapy. The sequences of the invention are useful as vaccines for inducing immunological response. The sequences of the invention are useful for treating cancers, infections, autoimmune disorders, haematopoietic disorders, wound healing disorders, cholesteryl ester storage disease, inflammation, congenital muscular dystrophy, junctional epidermolysis bullosa, Parkinson's disease, Huntington's chorea, multiple sclerosis, viral and bacterial infections, Alzheimer's disease, asthma, arthritis, allergies, schizophrenia, sbg44245FOA-associated disorders, septicemia, psoriasis, inflammatory bowel disease, transplant rejection, graft versus host disease, ischaemia, stroke, acute respiratory disease syndrome, restenosis, brain injury, AIDS, bone diseases, atherosclerosis, brain disorders including paraspranuclear palsy, myotonic dystrophy, depression, anxiety disorders and sleep disorders, cardiovascular diseases including congestive heart failure and myocardial infarction, respiratory diseases including chronic obstructive pulmonary disease, acute bronchitis and adult respiratory distress syndrome, liver disorders including hypercholesterolaemia, hypertriglyceridaemia, cirrhosis, viral and non-viral hepatitis, type II diabetes mellitus, renal disease including acute and chronic renal failure, glomerulonephritis, Fanconi's syndrome, cystinuria, skeletal muscle disorders including hypoglycaemia and tendinitis, gastrointestinal diseases including intestinal obstruction and tropical sprue, spleen disorders including hypersplenism, Hodgkin's disease and malignant lymphoma, testicular cancer, male reproductive diseases including low testosterone and male infertility. The present sequence is human pancreatic lipase gene.  
 XX Sequence 1383 BP; 399 A; 268 C; 302 G; 414 T; 0 other;  
 XX Query Match 87.9%; Score 1234.6; DB 24; Length 1383;  
 XX Best Local Similarity 93.8%; Pred. No. 0;  
 XX Matches 1314; Conservative 0; Mismatches 69; Indels 18; Gaps 2;  
 QY 1 ATGCTTGGAATTTGGATTGTTGCAATTTCTTGTCTTGGCACATCAAGAGAAAGAGTT 60  
 Db 1 ATGCTTGGAATTTGGATTGTTGCAATTTCTTGTCTTGGCACATCAAGAGAAAGAGTT 60  
 QY 61 TGTATGAAAGTTAGGTTGTTTCAAGAGTGGTTTACCATGGACAGGACTTTCTCAACA 120  
 Db 61 TGTATGAAAGTTAGGTTGTTTCAAGAGTGGTTTACCATGGACAGGACTTTCTCAACA 120  
 QY 121 GAGTTGGTAGGTTTACCTGCTCCAGAGAGTAACACTGGTTTCTGCTCTCACT 180  
 Db 121 GAGTTGGTAGGTTTACCTGCTCCAGAGAGTAACACTGGTTTCTGCTCTCACT 180  
 QY 181 ATACACATCCCAATGCCTATCAGGAGATCAGTGCCTTAATTTCTTCACTATCCAGGCC 240

```
181 ATACCAATCCCAATGCCATATCAGGAGATCAGTGGGTAAATCTTCAACTATCCAGCC 240
|||||
241 TCATATTTTGGAAACAGACAGATCACCAGTATCAACATAGCTGGATGGAAACAGATGGC 300
|||||
241 TCATATTTTGGAAACAGACAGATCACCAGTATCAACATAGCTGGATGGAAACAGATGGC 300
|||||
301 AAATGGCAGAGAGACATGTGCAATGTGTGTCAGCTGGAAGATATAAAATTCATTAAT 360
|||||
301 AAATGGCAGAGAGACATGTGCAATGTGTGTCAGCTGGAAGATATAAAATTCATTAAT 360
|||||
361 TTAGATTTGATCAACGGTTACGGGAATACATCCATCTGTAAACATCTCCGTGTGTT 420
|||||
361 TTAGATTTGATCAACGGTTACGGGAATACATCCATCTGTAAACATCTCCGTGTGTT 420
|||||
421 GGTGCTGAGTGGCTTATTTTATTGATGTTCTCATGAAAAAATTTGAATATCCCTTCT 480
|||||
421 GGTGCTGAGTGGCTTATTTTATTGATGTTCTCATGAAAAAATTTGAATATCCCTTCT 480
|||||
481 AAATGTCACATTTGATTTGGCCACAGCTTGGGACACACCTGGCTGGGAGCTGGGTCAAG 540
|||||
481 AAATGTCACATTTGATTTGGCCACAGCTTGGGACACACCTGGCTGGGAGCTGGGTCAAG 540
|||||
541 ATACGAGGCTTTGGAAGAAATAACTGGGTTGGACCCAGCTGGGCCATTTTCCACAACT 600
|||||
541 ATACGAGGCTTTGGAAGAAATAACTGGGTTGGACCCAGCTGGGCCATTTTCCACAACT 600
|||||
601 CCAAGGAGTCAAGCTAGACCCCTCGGATGCCAACTTTGTTGACGTTATTCATAAAT 660
|||||
601 CCAAGGAGTCAAGCTAGACCCCTCGGATGCCAACTTTGTTGACGTTATTCATAAAT 660
|||||
661 GCAGCTCGCATCTCTTTGAGCTTGTGTTGGAAACCATTTGATGTTGTTGATCTTTGAC 720
|||||
661 GCAGCTCGCATCTCTTTGAGCTTGTGTTGGAAACCATTTGATGTTGTTGATCTTTGAC 720
|||||
721 TTTTACCAATGGAGGAACACATCCAGATGTGAAGACTTAATACACCTTTACTG 780
|||||
721 TTTTACCAATGGAGGAACACATCCAGATGTGAAGACTTAATACACCTTTACTG 780
|||||
781 AAATTTAACTTCAATGCTTACAAAAGAAATGCTTCTTCTGACTGTACCATGCC 840
|||||
781 AAATTTAACTTCAATGCTTACAAAAGAAATGCTTCTTCTGACTGTACCATGCC 840
|||||
841 CGAAGTTATCAATTTTATGCTGAAAGCATTTTAACTCTGATGCAITTTATGCTTATCT 900
|||||
841 CGAAGTTATCAATTTTATGCTGAAAGCATTTTAACTCTGATGCAITTTATGCTTATCT 900
|||||
901 TGTAGATCTTACATCTTTTAAAGCAGAAATGCTTTCTTTTCCAAAGAGGTGTC 960
|||||
901 TGTAGATCTTACATCTTTTAAAGCAGAAATGCTTTCTTTTCCAAAGAGGTGTC 960
|||||
961 CCAACAAATGGGTCAATTTTGTCTGATAGATTTTCACTTCAAAAATATGAAGACTAATGATCA 1020
|||||
961 CCAACAAATGGGTCAATTTTGTCTGATAGATTTTCACTTCAAAAATATGAAGACTAATGATCA 1020
|||||
994 GTTACATAGGATAGATAGATAGAGATCATCTTCCATGTTTAAACCTTTCTCT 1013
|||||
1021 CATTATTTTAAACACAGGTGCTTTTCCCATTTCCCGTTGGAGGCAAAATGCTT 1080
|||||
1014 CCGTTTCTCTT-----CTTGTTTCTTATATCTAGGTTGGAGGACAAATGCTT 1062
|||||
1081 GTTAAACTCAGTGGAGGAGTCACTCAAGGACTGCTTTCTTCTGTTAGGCGGGCA 1140
|||||
1063 GTTAACTCAGTGGAGGAGTCACTCAAGGACTGCTTTCTTCTGTTAGGCGGGCA 1122
|||||
1141 ATTGGAAACTGGGAGTTTGGCATTGTCTAGTGGAAAACTTGAGCCAGGCACTTAC 1200
|||||
1123 GTTAGAAAACTGGGAGTTTGGCATTGTCTAGTGGAAAACTTGAGCCAGGCACTTAC 1182
|||||
1201 ACAAAATTAATGATGAGATGTTAAAGTTGGAAACATTACAGTGTTCAGTTCACTGG 1260
|||||
1183 ACAAAATTAATGATGAGATGTTAAAGTTGGAAACATTACAGTGTTCAGTTCACTGG 1242
|||||
1261 AAAAAACATTTTGTGAAGATTTCTCAGAAATAGTTGGGAGCAGAAATGGTATAAATACA 1320
|||||
```

1243 AAAAAACATTTTGTGAAGATTTCTCAGAAATAGTTGGGAGCAGAAATGGTATAAATACA 1302

1321 TCTGGAAATATGATATAAATCTACCTTCTGTAGCAAGACATTTATGGACCTAATTT 1380

1303 TCTGGAAATATGATATAAATCTACCTTCTGTAGCAAGACATTTATGGACCTAATTT 1362

1391 CTCAGAACCTGAAACCATGC 1401

1363 CTCAGAACCTGAAACCATGC 1383

RESULT 5

ABL60540

ID ABL60540 standard; cDNA; 1526 BP.

XX ABL60540;

AC

XX 27-AUG-2002 (first entry)

DT

XX Human lipid metabolism enzyme (LME)-4 cDNA (clone id: 7482937CB1).

DE

XX Human; lipid metabolism enzyme; LME; cytostatic; neuroprotective; gene;

KW nontropic; cerebroprotective; antiparkinsonian; antialzheimers; vaccine;

KW antileukemic; antimicrobial; anti-AIDS; cardiovascular; antiangiinal;

KW gene therapy; protein therapy; ss.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

XX CDS 259..1365

XX FT /\*tag= a

XX FT /product= "lipid metabolism enzyme"

XX FT /note= "LME-4"

XX

XX WO200229036-A2.

XX

XX 11-APR-2002.

XX

XX 05-OCT-2001; 2001WO-US31302.

XX

XX 06-OCT-2000; 2000US-238388P.

XX

XX 13-OCT-2000; 2000US-240616P.

XX

XX 02-NOV-2000; 2000US-245719P.

XX

XX 08-NOV-2000; 2000US-247503P.

XX

XX 17-NOV-2000; 2000US-249503P.

XX

XX (INCY-) INCYTE GENOMICS INC.

XX

XX Harland L, Arvizu C, Das D, Griffin JA, Baughn MR, Ding L;

XX

XX Walia NK, Yao MG, Lu Y, Elliott VS, Thangavelu K, Rankumar J;

XX

XX Lal PG, Tribouley CM;

XX

XX WPI: 2002-315862/35.

XX

XX P-PSDB; ABB08004.

XX

XX Lipid Metabolism Enzymes and nucleic acids, useful for preventing,

XX

XX diagnosing and treating e.g. cancer, Alzheimer's disease and

XX

XX Creutzfeld-Jakob disease -

XX

XX Claim 5; Page 123; 127pp; English.

XX

XX The invention relates to human lipid metabolism enzymes (LMEs) and

XX

XX encoding polynucleotides. The LMEs can be expressed by standard

XX

XX recombinant technology. The LME polypeptides, polynucleotides and

XX

XX modulators may be used in the prevention, diagnosis and treatment of

XX

XX diseases associated with inappropriate LME expression such as cancer

XX

XX (e.g. myeloma, sarcoma and breast cancer), neurological disorders (e.g.

XX

XX Parkinson's, Alzheimer's and multiple sclerosis), microbial infections

XX

XX (e.g. Creutzfeld-Jakob disease and Acquired Immune deficiency syndrome

XX

XX (AIDS)) and/or cardiovascular disorders (e.g. cardiomyopathy, angina

XX

XX pectoris and mitral valve prolapse). The present sequence represents the

XX

XX human LME-4 encoding cDNA.



Sequence 1526 BP; 440 A; 303 C; 325 G; 458 T; 0 other;

ry Match 79.5%; Score 1115.8; DB 24; Length 1526;  
t Local Similarity 94.3%; Pred. No. 4.8e-305;  
ches 1198; Conservative 0; Mismatches 2; Indels 71; Gaps 1;

205 GAGATCAGTGGCGTTAAATTCCTCAACTATCCAGGCTCATATTTGGACAGACAGATC 264  
|||||  
95 GAGATCAGTGGCGTTAAATTCCTCAACTATCCAGGCTCATATTTGGACAGACAGATC 154  
|||||  
265 ACCGATATCAACATAGCTGGATGGAAACAGATGCAAAATGGCAGAGACATGTCGAAT 324  
|||||  
155 ACCGATATCAACATAGCTGGATGGAAACAGATGCAAAATGGCAGAGACATGTCGAAT 214  
|||||  
325 ----- 324  
215 GTATGACATGAATAAGCTCCTTTTACACTAGCATGCGAGCTTTATGTTTAACTGAATG 274  
325 -----GTGTTGCTACAGCTGGAGATATATAATTCATTAAATTTAGATTGGATCA 373  
|||||  
275 TACTTTGCAAGGTGTTGCTACAGCTGGAGATATATAATTCATTAAATTTAGATTGGATCA 334  
|||||  
374 ACGGTTACCGGAATACATCCATGCTGTAACAATCTCCGTTGTTGGTGTGCTGAGGTGG 433  
|||||  
335 ACGGTTACCGGGATACATCCATGCTGTAACAATCTCCGTTGTTGGTGTGCTGAGGTGG 394  
|||||  
434 CTTATTTATGATGTTCTCATGAAATAATTTGAATATTCCTCTTAAAGTGCACCTTGA 493  
|||||  
395 CTTATTTATGATGTTCTCATGAAATAATTTGAATATTCCTCTTAAAGTGCACCTTGA 454  
|||||  
494 TTGGCCACAGCTTTGGGAGCACACCTCGGCTGGGAAAGCTGGCTCAAGATACACAGGCTTG 553  
|||||  
455 TTGGCCACAGCTTTGGGAGCACACCTCGGCTGGGAGCTGGCTCAAGATACACAGGCTTG 514  
|||||  
554 GAAGATAACTGGTTGAGCCAGCTGGGCAATTTTCCACAAACATCTCCAAAGGAATCA 613  
|||||  
515 GAAGATAACTGGTTGAGCCAGCTGGGCAATTTTCCACAAACATCTCCAAAGGAATCA 574  
|||||  
614 GGCTAGACCCCTCGGATGCGCACTTTGTTGAGCTTTATTCATACAAATGCGAGTCGCAATC 673  
|||||  
575 GGCTAGACCCCTCGGATGCGCACTTTGTTGAGCTTTATTCATACAAATGCGAGTCGCAATC 634  
|||||  
674 TCTTTGAGCTTGGTGTGGAACCAATGATGTTGGTGCATCTTTGACTTTTACCCAAATG 733  
|||||  
635 TCTTTGAGCTTGGTGTGGAACCAATGATGTTGGTGCATCTTTGACTTTTACCCAAATG 694  
|||||  
734 GAGGGAAGCACATGCCAGATGGAAGCTTAATTTACACCTTTTACTGAAATTTTAACTTCA 793  
|||||  
695 GAGGGAAGCACATGCCAGATGGAAGCTTAATTTACACCTTTTACTGAAATTTTAACTTCA 754  
|||||  
794 ATGCTTTACAAAAAGAAATGGCTTCTTTGACTGTAACCATGCCCGAAGTTATCAAT 853  
|||||  
755 ATGCTTTACAAAAAGAAATGGCTTCTTTGACTGTAACCATGCCCGAAGTTATCAAT 814  
|||||  
854 TTTATGCTGAAAGCAATCTTAATCTCGATGCAATTTATGCTTATCCTTTAGATCTTACA 913  
|||||  
815 TTTATGCTGAAAGCAATCTTAATCTCGATGCAATTTATGCTTATCCTTTAGATCTTACA 874  
|||||  
914 CATCTTTTAAAGCAGAAATTCCTTCTTTGTCACAAAGAGGTTGCCAACATGGGTC 973  
|||||  
875 CATCTTTTAAAGCAGAAATTCCTTCTTTGTCACAAAGAGGTTGCCAACATGGGTC 934  
|||||  
974 ATTTGCTGATAGATTTCACTTTCAAAAAATATGAAGCAATATGGAATCAATTTTAA 1033  
|||||  
935 ATTTGCTGATAGATTTCACTTTCAAAAAATATGAAGCAATATGGAATCAATTTTAA 994  
|||||  
1034 ACACAGGTCCTTTCCCATTTGCTTTGTCGAGGCACAAATTCCTGTTAACTCAGTG 1093  
|||||  
995 ACACAGGTCCTTTCCCATTTGCTTTGTCGAGGCACAAATTCCTGTTAACTCAGTG 1054  
|||||  
1094 GAAGCGAAGTCACTCAAGGAATGCTTTCTTTGTCGAGGCGGCAATTTGGGAAACATG 1153  
|||||  
1055 GAAGCGAAGTCACTCAAGGAATGCTTTCTTTGTCGAGGCGGCAATTTGGGAAACATG 1114  
|||||

QY 1154 GGGAGTTGGCCATTGTCAGTGGAAAACTTGAGCCAGGCAATGACATTACAAAAATTAATCG 1213  
Db 1115 GGGAGTTGGCCATTGTCAGTGGAAAACTTGAGCCAGGCAATGACATTACAAAAATTAATCG 1174  
QY 1214 ATGCAGATGTTAAAGTTGGAAACATTTACAGTGTTCAGTTCACTTGGAAAAACATTTGT 1273  
Db 1175 ATGCAGATGTTAAAGTTGGAAACATTTACAGTGTTCAGTTCACTTGGAAAAACATTTGT 1234  
QY 1274 TTGAAGATTTCTCAGAAATAAGTTGGGAGCAGAAAAATGGTCAATAATACATCTGGGAAATATG 1333  
Db 1235 TTGAAGATTTCTCAGAAATAAGTTGGGAGCAGAAAAATGGTCAATAATACATCTGGGAAATATG 1294  
QY 1334 GATATAAATCTACCTTCTGTAGCCAGACATTTATGGGACCTATATTTCTCAGAACCTGA 1393  
Db 1295 GATATAAATCTACCTTCTGTAGCCAGACATTTATGGGACCTATATTTCTCAGAACCTGA 1354  
QY 1394 AACCATGCTAA 1404  
Db 1355 AACCATGCTAA 1365  
RESULT 6  
ABA90354  
ID ABA90354 standard; cDNA; 1389 BP.  
XX  
AC ABA90354;  
XX  
DT 12-FEB-2002 (first entry)  
XX  
DE Human polynucleotide #29.  
XX  
KW Human; nootropic; neuroprotective; anticonvulsant; antidepressant;  
KW neuroleptic; tranquilizer; antiarrhythmic; cardiac; antischismatic;  
KW antinflammatory; antilipemic; hepatotropic; virucide; antidiabetic;  
KW nephrotoxic; anorectic; cytostatic; vaccine; neurological disease;  
KW cardiovascular disease; respiratory disease; liver disease;  
KW renal disease; skeletal muscle disease; gastrointestinal disease;  
KW placental disease; testicular cancer; male fertility; pancreatic disease;  
XX  
OS Homo sapiens.  
XX  
PN WO200181363-A1.  
XX  
PD 01-NOV-2001.  
XX  
PF 26-APR-2001; 2001WO-US13360.  
XX  
PR 27-APR-2000; 2000US-199963P.  
PR 11-MAY-2000; 2000US-203336P.  
PR 25-MAY-2000; 2000US-207087P.  
PR 26-MAY-2000; 2000US-207546P.  
XX  
(SMIK ) SMITHKLINE BEECHAM CORP.  
PA (SMIK ) SMITHKLINE BEECHAM PLC.  
XX  
PI Agarwal P, Murdock PR, Rizvi SK, Smith RF, Xiang Z, Kabnick KS;  
PI Lai Y, Xie Q;  
XX  
DR WPI; 2002-041392/05.  
DR P-PSDB; ABB53289.  
XX  
PT Novel polypeptides and polynucleotides useful as a vaccine for  
PT preventing and treating diseases associated the polypeptide, e.g.  
PT Alzheimer's disease, dyslipidemia, obesity, diabetes, infertility,  
PT asthma, amnesia  
XX  
PS Claim 2; Page 60-61; 116pp; English.  
XX  
CC The invention relates to an isolated polypeptide comprising a 277, 480,  
CC 583, 581, 628, 428, 638, 229, 310, 841, 241, 369, 382, 185, 586, 1026,  
CC 844, 782, 262, 394, 471, 485, 286, 533, 495, 350, 619, 490, 462, 255,

784, 252, 593, 472, 607, 781, 640, 686 or 154 amino acid sequence as given in the specification. The polypeptides, modulators of the polypeptides and antibodies against the polypeptides are useful for treating Alzheimer's, parasupranuclear palsy, Huntington's disease, myotonic dystrophy, anorexia and depression; cardiovascular diseases including congestive heart failure, Hodgkin's disease and myocardial infarction; respiratory diseases including asthma, chronic obstructive pulmonary disease, cystic fibrosis and adult respiratory distress syndrome; liver diseases including hypercholesterolemia, cirrhosis, viral and nonviral hepatitis, Type II diabetes mellitus, and impaired glucose tolerance; renal disease including renal failure, acute tubular necrosis and glomerulonephritis; skeletal muscle diseases including Eulenburg's disease, hypoglycemia and obesity; gastrointestinal diseases including myotonia congenita and intestinal obstruction; lymph diseases including lymphaglectasia; diseases of placenta including choriocarcinoma; diseases of testes including testicular cancer, male reproductive diseases including low testosterone and male infertility; and disease of pancreas including diabetic ketoacidosis, Type 1 and 2 diabetes and obesity. The present genetic encodes a polypeptide of the invention.

Sequence 1389 BP; 392 A; 294 C; 304 G; 399 T; 0 other;  
ery Match 69.7%; Score 978.8; DB 24; Length 1389;  
st Local Similarity 82.6%; Pred. No. 2.6e-266; Indels 12; Gaps 1;  
ches 1137; Conservative 0; Mismatches 227;  
29 TGTCTTTGGGCATCAAGAGAAAGAGTTCTCTATGAAGTTAGGGTGTTCAAAG 88  
26 TGCTGCTGGGAGCAGTAGCAGGAAAGAGTTTCTACGAAAGACTCGGCTGCTCAGTG 85  
89 ATGGTTTACCATGGACAGGACTTCTCACAGAGTTGGTAGTTTACCTGCTCCAG 148  
86 ATGACTCCCATGGTTCAGGAATTACGGAAGACCCCTCCATATTTGCTGTGCTCCAA 145  
149 AGAAGATAAACACTGGTTCTGCTCTACACTATACATAATCCCAATGCTATCAGGAGA 208  
146 AAGATGTCAACACCCGCTCTCTCTATATACTAATGAGAACCCCAACAACTTCAAGAGA 205  
209 TCAGTGGGTTAATTTCTCACTATCCAGCTCATATTTTGGACAGACAAGATCACCC 268  
206 TCAGTGGGTTAATTTCTCACTATCCAGCTCATATTTTGGACAGACAAGATCACCC 265  
269 GTATCAACATAGCTGGATGGAAGAACAGATGCGCAATGCGCAGAGACATGTGCAATGTGT 328  
266 GTATCAACATAGCTGGATGGAAGAACAGATGCGCAATGCGCAGAGACATGTGCAATGTGT 325  
329 TGCTACAGCTGGAAGATATAATTGCTAATATTAGATTGGATCAACGGTTACGGGAAT 388  
326 TGCTACAGCTGGAAGATATAATTGCTAATATTAGATTGGATCAACGGTTACGGGAAT 385  
389 ACATCCATGCTCTAAACAATCTCGTGTGTTGTTGCTGAGTGGCTTATTTATTGATG 448  
386 ACATCCATGCTCTAAACAATCTCGTGTGTTGTTGCTGAGTGGCTTATTTATTGATG 445  
449 TTCTCATGAAAAAATTTGAATATTTCCCTCTAAAGTGCACCTGATTTGGCCACAGCTTG 508  
446 TTCTCATGAAAAAATTTGAATATTTCCCTCTAAAGTGCACCTGATTTGGCCACAGCTTG 505  
509 GAGCACCTGCTGGGAGCTGGTCAAGATACAGGCTTGGAGATCACTTGGGTTGGGT 568  
506 GAGCACCTGCTGGGAGCTGGTCAAGATACAGGCTTGGAGATCACTTGGGTTGGGT 565  
569 TGGACCCAGCTGGGCTCAATTTTCCACACACTCCCAAGGAAGTCCAGCTAGACCCCTCG 628  
566 AGCATGCCCTGCTGAGTTGGGCTGGAGTGTGCAACCGAGGGCTACTGTTATCTGCTACTC 625  
629 ATGCCAATTTGTTGACGCTTATTCATAAATGAGCTCGATCTCTTTGAGCTTTGGTG 688  
626 TGGTAACAATGCTCAACTTCGTAGTACAAATCACTGATGCAACTCCCATATAATCCCC 685  
689 TTGGAACCATTTGATGCTGTGCTCATCTTGACTTTTACCCAAATGGAGGGAAGACATGC 748

Db AGTGGATGAGGGGCACTTCGGGAACCCAGTAACCCCTCTCTGTACTCTCTCTGTGCT 745  
Qy CAGGATGTGAGACTTAATTACACCTTTACTGAAATTTAACTTCAATGCTTACAAAAAG 808  
Db TATGGTTGGCTGATCTGGCTCGTCTCACTCGTATGTTTGGGCT-----G 793  
Qy AAATGGCTTCTCTTTGACTGTAAACCATGCCGAGTTATCAATTTTATGCTGAAAGCA 868  
Db AAATGGCTTCTCTTTGACTGTAAACCATGCCGAGTTATCAATTTTATGCTGAAAGCA 853  
Qy TTCTAATCTGATGATTTATTTGCTTATCTTGTAGATCTTACACATCTTTTAAAGGAG 928  
Db TTCTAATCTGATGATTTATTTGCTTATCTTGTAGATCTTACACATCTTTTAAAGGAG 913  
Qy GAAATGCTTCTTTTGTTCAAAGAGAGTTGCCCCAATCGGTGATTTTGTGCTAGAT 988  
Db GAAATGCTTCTTTTGTTCAAAGAGAGTTGCCCCAATCGGTGATTTTGTGCTAGAT 973  
Qy TTCACTTCAAAAATATGAAGACTTAATGGATCACTATTTTAAACACAGGGTCCCTTT 1048  
Db TTCACTTCAAAAATATGAAGACTTAATGGATCACTATTTTAAACACAGGGTCCCTTT 1033  
Qy CCCATTGCCCCGTTGGAGGCACAAATTTGTTTAACTCAGTGGAGCGAAGTCACTC 1108  
Db CCCATTGCCCCGTTGGAGGCACAAATTTGTTTAACTCAGTGGAGCGAAGTCACTC 1093  
Qy AAGAACTGCTTCTTCTGTTAGGGGGGAGTTAGGAAAACTGGGGAGTTTGGCAATG 1153  
Db AAGAACTGCTTCTTCTGTTAGGGGGGAGTTAGGAAAACTGGGGAGTTTGGCAATG 1153  
Qy TCAGTGGAAACTTGGAGCGATGATCACTACACAAATTAATTCGATCGAGATGTTAAG 1228  
Db TCAGTGGAAACTTGGAGCGATGATCACTACACAAATTAATTCGATCGAGATGTTAAG 1213  
Qy TTGAAACATTACAAGTGTTCAGTTTCATCTGGAAGAAACATTTGTTGAAGATTTCTAGA 1288  
Db TTGAAACATTACAAGTGTTCAGTTTCATCTGGAAGAAACATTTGTTGAAGATTTCTAGA 1273  
Qy ATAGTTTGGAGCAGAAATGGTGAATAATACATCTGGGAAATATGGATATAATCTACT 1348  
Db ATAGTTTGGAGCAGAAATGGTGAATAATACATCTGGGAAATATGGATATAATCTACT 1333  
Qy TCTGTAGCCCAAGACATTTATGGGACCTTAATTTCTCCAGAACCTGAAACCATGCTAA 1404  
Db TCTGTAGCCCAAGACATTTATGGGACCTTAATTTCTCCAGAACCTGAAACCATGCTAA 1389

RESULT 7  
ABN60002  
ID ABN60002 standard; cDNA; 1540 BP.  
XX  
AC ABN60002;  
XX  
DT 28-JUN-2002 (first entry)  
XX  
DE Novel human coding sequence SEQ ID NO: 413.  
XX  
KW Human; antianaemic; vulnery; antiinflammatory; immunomodulator;  
KW antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;  
KW neuroprotective; antiparkinsonian; protein therapy; EST;  
XX expressed sequence tag; gene; ss.  
OS Homo sapiens.  
XX  
PN WO200222660-A2.  
XX  
PD 21-MAR-2002.  
XX  
PF 10-SEP-2001; 2001WO-US26015.  
XX  
PR 11-SEP-2000; 2000US-0659671.  
XX

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren P;  
Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
WPI; 2002-292408/33.  
P-PSDB; ABB97589.

An isolated polynucleotide for treating diseases associated with its  
encoded polypeptide such as cancer and multiple sclerosis -

Claim 1; SEQ ID NO 413; 509pp; English.

The present invention provides the protein and coding sequences of 444  
novel human proteins. These were isolated from expressed sequences tags  
(ESTs). They can be used to stimulate cell growth, to regulate  
haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
e.g. in burn treatment, to regulate the immune system e.g. to treat  
multiple sclerosis, to regulate activin or inhibin e.g. to treat  
infertility, to regulate haemostasis or thrombolysis e.g. to treat  
stroke and cancer, to screen for drugs, to treat inflammatory conditions  
e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.  
Parkinson's disease. The present sequence is a coding sequence of the  
invention.

Sequence 1540 BP; 477 A; 276 C; 324 G; 463 T; 0 other;

ry Match 57.5%; Score 808; DB 24; Length 1540;  
t Local Similarity 100.0%; Pred. No. 5.8e-218;  
ches 808; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCTTGGAAATTGGATTGCTTCATCTTGTCTTGGCATCATCAAGGAAAGAGATT 60  
86 ATGCTTGGAAATTGGATTGCTTCATCTTGTCTTGGCATCATCAAGGAAAGAGATT 145  
61 TGCATGAAGGTTAGGGTGTTCAAAGATGGTTTACCATGGACGAGACATCTTCAACA 120  
146 TGCATGAAGGTTAGGGTGTTCAAAGATGGTTTACCATGGACGAGACATCTTCAACA 205  
121 GAGTGGTGGTTTACCTGGTCTCAGAGAAGATAAACAATCTGTTCTGCTCTCAACT 180  
206 GAGTGGTGGTTTACCTGGTCTCAGAGAAGATAAACAATCTGTTCTGCTCTCAACT 265  
181 ATACAAATCCCAATGCCATCAGAGATCAGTGGGTTAATCTTCAACTATCAAGGCC 240  
266 ATACAAATCCCAATGCCATCAGAGATCAGTGGGTTAATCTTCAACTATCAAGGCC 325  
241 TCATATTTGGAAACAGACAAAGATACCCGTTATCAACATAGCTGGATGGAACAGATGGC 300  
326 TCATATTTGGAAACAGACAAAGATACCCGTTATCAACATAGCTGGATGGAACAGATGGC 385  
301 AATGCGCAGAGACATGTGCAATGTTGCTACAGCTGGAAGATATAATTCGATTAAT 360  
386 AATGCGCAGAGACATGTGCAATGTTGCTACAGCTGGAAGATATAATTCGATTAAT 445  
361 TTAGATTGGATCAACGGTTTACGGGAATACATCCATGCTGTAAACAATCTCCGTTGTT 420  
446 TTAGATTGGATCAACGGTTTACGGGAATACATCCATGCTGTAAACAATCTCCGTTGTT 505  
421 GGTGCTGAGTGGCTTATTTATGATGTTCTCATGAAAAAATTTGAATATCCCTTCT 480  
506 GGTGCTGAGTGGCTTATTTATGATGTTCTCATGAAAAAATTTGAATATCCCTTCT 565  
481 AAAGTGCATCTGATTGGCCACAGCTGGGAGCACACCTGGCTGGGGAAGCTGGTCAAGG 540  
566 AAAGTGCATCTGATTGGCCACAGCTGGGAGCACACCTGGCTGGGGAAGCTGGTCAAGG 625  
541 ATACAGGCTTGGAAAGATAAATCTGGTTGGACCACTGGCTGGGCAATTTTCCACAACT 600  
626 ATACAGGCTTGGAAAGATAAATCTGGTTGGACCACTGGCTGGGCAATTTTCCACAACT 685  
601 CCAGAGAGTCAAGGCTAGACCCCTCGATGCGCACTTGTGACGTATTTCATCACAAT 660

Db 686 CCAAGGAAGTCAGGCTAGACCCCTCGATGCCAACTTGTGTGACGTTATTATCATCAAAAT 745  
Qy 661 GCAGCTCGCATCTCTTTGAGCTTGGTGTGGAACCATGATGCTTGTGGTCACTTGAC 720  
Db 746 GCAGCTCGCATCTCTTTGAGCTTGGTGTGGAACCATGATGCTTGTGGTCACTTGAC 805  
Qy 721 TTTTACCAAAATGGAGGAAGCACATGCCAGGATGTGAAGACTTTAATTACACCTTTACTG 780  
Db 806 TTTTACCAAAATGGAGGAAGCACATGCCAGGATGTGAAGACTTTAATTACACCTTTACTG 865  
Qy 781 AAATTAACTTCAATGCTTACAAAAG 808  
Db 866 AAATTAACTTCAATGCTTACAAAAG 893  
RESULT 8  
ID ABL69985 standard; DNA; 1481 BP.  
XX ABL69985;  
AC ABL69985;  
XX  
DT 15-MAY-2002 (first entry)  
DE Pancreas cancer related gene sequence SEQ ID NO:8322.  
XX  
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
gene; ds.  
XX Homo sapiens.  
OS  
XX  
XX WO200194629-A2.  
XX  
XX 13-DEC-2001.  
XX  
XX 30-MAY-2001; 2001WO-US10838.  
XX  
XX 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.  
PR 25-SEP-2000; 2000US-234923P.  
PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235134P.  
PR 25-SEP-2000; 2000US-235280P.  
PR 26-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 27-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 27-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236028P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 28-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236842P.  
PR 29-SEP-2000; 2000US-236891P.  
PR 02-OCT-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 02-OCT-2000; 2000US-237316P.

03-OCT-2000; 2000US-237425P.  
 03-OCT-2000; 2000US-237598P.  
 03-OCT-2000; 2000US-237604P.  
 03-OCT-2000; 2000US-237606P.  
 03-OCT-2000; 2000US-237608P.  
 01-NOV-2000; 2000US-244867P.  
 01-NOV-2000; 2000US-245084P.  
 (AVAL-) AVALON PHARM.  
 Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 Soppet DR, Weaver Z;  
 WPI; 2002-188264/24.  
 Screening for anti-neoplastic agent involves exposing cells to a  
 chemical agent to be tested for anti-neoplastic activity, and  
 determining a change in expression of a gene of a signature gene set -  
 Claim 1; SEQ ID 8322; 44pp; English.

The present invention describes a method (M1) for screening for an  
 anti-neoplastic agent. The method involves exposing cells to a chemical  
 agent to be tested for anti-neoplastic activity, determining a change in  
 expression of at least one gene (I) of a signature gene set, where (I)  
 comprises a sequence (S) selected from 8447 sequences (given in ABL6164  
 to ABL70110), or is at least 95% identical to (S), where a change in  
 expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 activity and can be used in gene therapy. M1 can be used for screening  
 an anti-neoplastic agent, and can be used for producing a product which  
 is the data collected with respect to the anti-neoplastic agent as a  
 result of M1, and the data is sufficient to convey the chemical  
 structure and/or properties of the agent. M1 can be used in the  
 treatment of cancer such as colon, breast, stomach, lung, thyroid,  
 esophageal, ovarian, kidney, prostate or pancreatic cancer,  
 adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
 infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
 carcinoma, papillary carcinoma and Wilms' tumour.

Sequence 1481 BP; 399 A; 371 C; 379 G; 332 T; 0 other;

ary Match 27.5%; Score 386.2; DB 24; Length 1481;  
 st Local Similarity 57.1%; Pred. No. 1.2e-98;  
 tches 807; Conservative 0; Mismatches 588; Indels 18; Gaps 5;

1 ATGCTTGGAATTCGATTTGGTCTTCTCTCTTTGGACATCAAGGAAAGAGTT 60  
 20 ATGCTGATCTCTGGCAATCACTTTCTCTGCTGGAGCAGCCAAAGAAAGAGTT 79  
 61 TGCTATGAAGAGTTAGGTTGTTTCAAGATGGTTTACCATGGACAGGACTTTCTCAACA 120  
 80 TGCTATGAGGACCTCGGTGCTTTCTGACACTGAGCCCTGGGGCGGGACAGCAATCAGG 139  
 121 GAGTTGGTGGTTTACCTGCTCTCCAGAGAGATAAAGTCTGTTCTGCTCTACACT 180  
 140 CCCCTGAAATTTCTCCCTGGAGCCCTGGAAGATCGGACCCGCTTCTGCTGTACACC 199  
 181 ATACACATCCCAATGCTCATCAGGAGATCAGTGGGTTAATTTCTCAACTATCAAGCC 240  
 200 AATGAAACCCCAACACTTTCATTTCTCTCTCTGATCCATCAACATTTGAGGCA 259  
 241 TCATATTTTGGAAACAGACATGATCCCGTATCAACATAGTGG-----ATGGAAACA 294  
 260 TCAATTTTCAATTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 319  
 295 GATGGCAATGGCAGAGAGATGATGCAATGTTGCTACAGCTGGAAGATATAAATGCG 354  
 320 GATGAGAGCTGGGTGACAGACATGTCGAAGAACTGTTGAGGTGAGGAGGTGAATGCG 379  
 355 ATTAATTTAGATTCGATCAACGGTTTACGGG---AATACATCATGCTGTAAACAACTCTC 411  
 380 ATCTCGGTGGAGCTGGAAGAGAGGCTTCCCAAGCCACCTACACACAGGCTGCCCAACAGTG 439

QY 412 CGTCTGTTGTTGCTGAGGTGGCTTATTTTATTGATGTTCTCATGAAAAAATTTGAATAT 471  
 DB 440 CGAGTGGTGGCGCCAGGTGGCCAGATGCTCGACATCTCTTTCAGAGATATAGCTAC 499  
 QY 472 TCCCTTTCTAAAGTGCATTTGATTTGGCCACAGCTTGGGAGACACCTGGCTGGGAAGCT 531  
 DB 500 CCCCTTTCCAAAGTTTCACTCTCATTTGGCCACAGCTTGGGAGCCACGCTGGCTGGAGAGCA 559  
 QY 532 GGGTCAAGGATACAGGCTTGGAGCAATAAATGCTGGTTGGACCCAGCTGGCCATTTTC 591  
 DB 560 GGAAGCAGACTCCAGGCTTGGAGCAGATTAAGGTTGGATCTCTGTAGAACAGATTTC 619  
 QY 592 CACAACACTCCAAAGGAAGTCAAGCTAGACCCCTCGGATGCCAACCTTTGTTGACGTTAT 651  
 DB 620 GAGAGTACTCTCGAAGAGGTGCGACTTGTATCCCTCTGATGCTGACTTTGTTGATGATT 679  
 QY 652 CATACAAATGCAGCTCGCATCTCTTTGAGCTTGGTGGTGGAAACCATTTGATGCTTGTGTT 711  
 DB 680 CACAGGATGAGCTCCCTGATCCCATTTCTTGGGTTTTGGAAACGAAACCAACAGATGGGT 739  
 QY 712 CATCTTGACTTTTACCCAAATGGAGGGAAGACATGCCAGGATGTGAAGACTTAATTACA 771  
 DB 740 CATCTTGACTTTCTCCCAATGGAGGAGAGATGCCGGATGCAAGAA---GAATGCC 796  
 QY 772 CCTTTACTGAAATTTAACTTCAATGCTTACAAAAGAAATGCTTCTCTTTGACTGT 831  
 DB 797 CTGTCTCAGATCTGATCTAGATGGCATCTGGGCGGGAACCCGGGACTTTGTGGCTTGC 856  
 QY 832 AACCATGCCCGAAGTTATCAATTTTATGCTGAAGCATTTCTTAATCTCTGATGACTTTATT 891  
 DB 857 AATCACCCTAAGAGCTTCAAGATTTACTTTGGAAGCATCTCTCAATCCCGATGGTTTGTCT 916  
 QY 892 GCTTATCTTTGATGCTTACACATCTTTTAAAGCAGGAAATGCTTCTTTGTTTCCAAA 951  
 DB 917 GCATATCCCTGCACTTCTTACAGTCTCTTTGAGTGTGACAGTCTTCCCGTGTCCAGAT 976  
 QY 952 GAAGGTTGGCCCAACATGGTCTATTTGCTGATGATTTTCACTTCAAAAATATGAAGACT 1011  
 DB 977 CAAGGATGCCACAGATGGTCTATGCTGATTAATTTGCTGGCAGGACAGTGA---A 1033  
 QY 1012 AATGGATCACTTTTAAACACAGGCTCCCTTTCCCATTTTGGCCGTTGGAGGCAC 1071  
 DB 1034 GAGCAGCAGAAATTTCTTTGAACACAGGAGAGGCTAGCAATTTCCGCTCGCTGGAGATAT 1093  
 QY 1072 AATTTCTGTTTAACTCAGTGGAGCGAGTCACTCAAGGACTGCTTCTTCTGCTGA 1131  
 DB 1094 GGGGTTTCAATCACTGTCTGAAAGAACAGCCACT---GGTCAAGTCAAAAGTTGCTTTG 1150  
 QY 1132 GCGGGGCAATTTGGGAAAACTGGGAGTTTGGCATTTGTCAAGTGGAAAACTTTGAGCCAGGC 1191  
 DB 1151 TTTGGAATAAGGGAACACTCACAGTACAGCATCTTTCAGGGGATTTCTCAAAACAGGC 1210  
 QY 1192 ATGACTTACAAAAATTAATCGATGAGATGTTACGTTGGAACATTTACAGTGTTCAG 1251  
 DB 1211 TCAACCCATCTCTATGAGTTTGTATGCAAGCTGGATGTTGGAACAATTTGAGAAAGTCAAG 1270  
 QY 1252 TTCACTCGAAAAAATTTGTTTGAAGATTTCTCAGAATAAGTTGGGAGCAGAAATGGTG 1311  
 DB 1271 TTTCTTTGGATTAACAATGATGATTAATCCACCCCTCCCAAGTGGGTGCCACCAAGATC 1330  
 QY 1312 ATAAATACATCTGGGAAATATGGAATATAAATCTACTCTCTGTAGCCAAAGACATTTATGGA 1371  
 DB 1331 ACTGTGCAAAAGGGGAG 1390  
 QY 1372 CCTAATATTTCTCAGAACCTGAAACCATGCTAA 1404  
 DB 1391 GAAGACACGCTGCTACCCCTCAGCCCTGCTTAA 1423

RESULT 9  
 AAS81082  
 ID AAS81082 standard; cDNA; 1427 BP.  
 XX



1248 TCAGTTCATCTGGAAAAACATTTGTTTGAAGATTCTCAGATAAAGTTGGGAGCAGAAAT 1307  
1271 CAAGTTTCTTTGGATAACAATGTGATAAATCCACCTCCCAAGCTGGGTGCCACCA 1330  
1308 GGTGATAATACATCTGGAAATATGGATATAAATCTACCTTCTGTAGCCAAAGCAATTAT 1367  
1331 GATCACTGTGCAAGAGGAGAGAGACACAGTGTACAACTTCTGTAGCGAAGACACAGT 1390  
1368 GGGACCTAATATCTCCAGAACCTTGAAACCATGCTAA 1404  
1391 GCGGGAAGACACGCTGCTCACCTCACGCCCTGCTAA 1427

LT 10

9552

ABL59552 standard; cDNA; 1454 BP.

ABL59552;

16-JUL-2002 (first entry)

Human triacylglycerol lipase cDNA SEQ ID NO:52.

Human; triacylglycerol lipase; enzyme; tumour; lipid associated gene; lipid metabolism; lipid synthesis; gene; ss.

Homo sapiens.

WO200227028-A1.

04-APR-2002.

27-SEP-2001; 2001WO-US030366.

28-SEP-2000; 2000US-0676052.

(ATAI-) ATAIRGIN TECHNOLOGIES INC.

Skinner MK, Patton JL, Chaudhary J;

WPI; 2002-402054/43.

Identifying tumor characteristics in a tissue sample taken from a patient, involves determining the copy number or expression level of genes associated with lipid metabolism, synthesis or action -

Example 1; Page 109-110; 113pp; English.

The present invention describes a method for identifying tumour characteristics, comprising measuring a copy number or expression level of at least two genes associated with lipid metabolism, synthesis, or action in cells from a patient tissue sample, and comparing the results with a copy number or expression level of the genes in a normal cell. Also described is an array of nucleic acid polymers immobilised on a solid support, comprising a solid support, at least two different nucleic acid polymers which are each specific for a different gene associated with lipid metabolism, synthesis or action, where each nucleic acid polymer is located at a predetermined position on the solid support, and the array comprises nucleic acid polymers which are specific for less than 100 genes other than the selected genes. The method is useful for determining tumour characteristics in a tissue sample taken from a patient. The present sequence represents a human lipid-associated gene related cDNA sequence, which is used in the exemplification of the present invention.

Sequence 1454 BP; 415 A; 313 C; 349 G; 377 T; 0 other;

ery Match 24.4%; Score 342.6; DB 24; Length 1454;  
st Local Similarity 57.0%; Pred. No. 2.6e-86;  
tches 791; Conservative 0; Mismatches 569; Indels 27; Gaps 8;

29 TGTCTTGGCACATCAAGAGCAAGAGTTTCTATGAAGTTAGGTTGTTTCAAG 88

Db 26 TGCTGTGGGAGCAGTAGCAGGAAAAAGAGTTTGTCTACGAAAGACTCGGCTGCTTCAGTG 85  
Qy 89 ATGCTTTACCATGGACGACGACTTTCTACACAGAGTTGGTAGGTTTACCTGGTCTCCAG 148  
Db 86 ATGACTCCCATCGTCTAGGAATTCGGAAGAGCCCTCCATATATTGGCTTGGTCTCCAA 145  
Qy 149 AGAAGATAAACACTCGTTTCTCTCTACACTATACAAATCCCAATGCTTATCAGGAGA 208  
Db 146 AAGATGCAACACCCGCTCTCTCTATATATACTAATAGAAACCCAAACAATTTCAAGAAG 205  
Qy 209 TCAGTGGGTTAATTTCTTCAACTATCCAGCCTCATATTTTGGAAACAGACAAGATCACCC 268  
Db 206 T---TGCCGCGAGATTTCATCAAGCATCAGTGGCTCCAAATTTCAAAAACAAATAGAAAACATC 262  
Qy 269 GTATCAACATAGCTGGAT-----GGAACACAGATGGCAATGGCAGAGAGACATGTGCA 322  
Db 263 GCITTTATTTTCATGGATTATAGACAAGGAGAGAAAACCTGGCTGGCCATGTGTGCA 322  
Qy 323 ATGTGTTGCTACAGCTGGAAAGATATAAATTTGCAATTAATTTAGATGGATCAACGGTTCA 382  
Db 323 AGAATCTGTTCAAGGTGGAAAGTGTGAACCTGTATCTGTGGACTGGAAAGGTGGCTCCC 382  
Qy 383 GG---GAATACATCCATGCTGTAAACAATCTCGTGTCTGTCTGCTGAGGTGGCTTATT 439  
Db 383 GAATGATACACACAGCCTCCAGAACATCAGGATCTGTGGAGCAGAGAGTGGCATATT 442  
Qy 440 TTATTGATGTTCTCATGAAAAAATTTGAAATATTTCCCTTTCTAAAGTGCACTTGAATGGCC 499  
Db 443 TTGTTGAATTTCTTCAGTGGCGCTTCGGTTACTCACCTTCCAACGTGCATGTCAITGGCC 502  
Qy 500 ACAGCTTGGGAGCACACCTGGCTGGGGAGCTGGGTCAAGGATACCAGG---CCTTGGAA 556  
Db 503 ACAGCTGGGTGCCAGCTGCTGGGGAGCTGGAGGAGAACCAATGGGACCATTTGGAC 562  
Qy 557 GAATAACTGGGTTGGACCCAGCTGGGCCATTTTTCACAAACACTCCAAAGGAAGTCAGGC 616  
Db 553 GCATCACAGGTTGGACCCAGCAGAACCTTGTCTTCAGGCGACACCTGAATTTAGTCCGAT 622  
Qy 617 TAGACCTCCGATGCCAATTTGTTGAGCTTATTCATACAAATGACAGCTCGCATCTCT 676  
Db 623 TGGACCCCGCAGTGCCTGCTGGGATGTAATTTGCGGATGTAATTCACACGGATGGTGGCCCTAGTCC 682  
Qy 677 TTGAGCTGTGTTGGAACCATTTGATGCTTGTGGTTCATCTTGACATTTTACCCAAATGGAG 736  
Db 683 CCAATTTGGGTTTGGAAATGAGCCAACTGCTGGGCCACCTAGATTCTTCCAAATGGAG 742  
Qy 737 GGAAGCACATGCCAGGATGGAAGACTTAATTTACACCTTTACTGAAATTTAACTTCAATG 796  
Db 743 GAGTGGAAATGCTCGTGTGTAATAAAGAACATTT---CTCTCTCAGATTGTGGACATAGACG 799  
Qy 797 CTTACAAAAAAGAAATGGCTTCTCTTGTGACTGTAAACCATGCCGAAAGTTATCAATTTT 856  
Db 800 GAATCTGGGAAGGACTCGAGACTTTGGCGCTGTAAATCACTTAAGAACCTACAAATATT 859  
Qy 857 ATGCTGAAGACATTTAAATCCTGATGCAATTTATTGCTTATCTCTTGTAGATCCTACAT 916  
Db 860 ACATGATAGACTCGTCAACCTGATGGCTTTGCTGGATTCCCTGTGCTCTTACAAAG 919  
Qy 917 CTTTAAAGCAGAAATTTGCTTTTGTTCGAAAGAGTTGCCCAACATGGGTCATT 976  
Db 920 TCTTCACTGCAACCAAGTGTTTCCCTTGTCCAAAGTGGAGCTGCCACAGATGGGTCACT 979  
Qy 977 TTGCTGATGATTTCACTTCAAAAATATGAAGACTAAATGGAATCACAATTTTAAAAACA 1036  
Db 980 ATGCTGATGATATCTCTGGGAAA---ACAAATGATGTGGCCAGAAAATTTTATCTAGACA 1036  
Qy 1037 CAGGTCCTTTTCCCATTTGCCGTTGGAGCACAATTTGCTGTAAACTCAGTGGNA 1096  
Db 1037 CTGGTGAATGCCATAAATTTGCACGTTGGAGGTATAAGGTATCTGTGCACACTGTCTGGAA 1096  
Qy 1097 GCGAAGTCACTCAAGGAACTGTCTTTCTTCTGTTAGGCGGGGCAATTTGGAAAAACTGGGG 1156  
Db 1097 AAAAGTTAC---AGGACACATACATAGTTTCTTTTGTTCGAAATTAAGGAACTCTNAGC 1153



1157 AGTTTGCCATTGTCAGTGGAAACCTTGAGCCAGGCATGACTTACACAAAATTAATCGATG 1216  
 |||||  
 1154 AGTATGAAATTTTCAAGGGCACTCTCAAAACAGATAGTACTCATTCCTCAATGAATTTGACT 1213  
 |||||  
 1217 CAGATGTTAACTGGAACATTAACAAGTGTTCAGTTTCATCTCGAATAAACATTTGTTG 1276  
 |||||  
 1214 CAGATGTGATGTTTGGGACCTTGCAGATGGTTAAATTTATTTGGTATAACAATGTGATCA 1273  
 |||||  
 1277 AAGATTCTCAGAAATAAGTTGGAGCAGAAATGGTGATAAATACATCTCGGGAATAATGAT 1336  
 |||||  
 1274 ACCCACTTTACTAGATGGAGGATCCAGATTATAGTGGAGAC--AAATGTTGGAA 1330  
 |||||  
 1337 ATAAATCTACCTTCTGTAGCCAGACATATATGGACCTTAATTTCTCCAGAACCTGAAC 1396  
 |||||  
 1331 AACAGTTCAACTTCTGTAGTCCAGAAACCGTCAGGAGGAAGTTCTGTCTACCCCTCAC 1390  
 |||||  
 1397 CATGCTA 1403  
 |||||  
 1391 CQTGTTA 1397  
 |||||

r 11

205

ABK44205 standard; cDNA; 1471 BP.

ABK44205;

21-MAY-2002 (first entry)

cDNA #145 encoding human pancreatic tumour protein #2.

human; pancreatic tumour protein; immune response; pancreatic cancer;  
 development of cancer; cancer progression; cytostatic; gene; ss.

Homo sapiens.

NO200212331-A2.

14-FEB-2002.

36-AUG-2001; 2001WO-US24619.

37-AUG-2000; 2000US-223130P.

30-JAN-2001; 2001US-265447P.

15-MAY-2001; 2001US-291201P.

(CORI-) CORIXA CORP.

Pyle RA, Xu J, Kalos MD;

NPI; 2002-241741/29.

P-PSDB; AAU87690.

Novel polynucleotide encoding pancreatic tumour polypeptides, useful in  
 pharmaceutical compositions, e.g. vaccines, for treating pancreatic  
 cancers -

Claim 1; Page 157; 167pp; English.

The present invention relates to the isolation of cDNA sequences  
 encoding human pancreatic tumour proteins. The polynucleotide  
 sequences encoding human pancreatic tumour proteins are useful for  
 stimulating an immune response in a patient and treating pancreatic  
 cancer in a patient. A host cell that expresses these polynucleotides  
 is useful for determining the presence of cancer in a patient. A  
 composition comprising the polynucleotide, its encoded protein, or an  
 antibody that binds to the protein may be used in the diagnosis,  
 prevention and/or treatment of diseases, particularly pancreatic  
 cancer. The sequences of the invention are also useful in pharmaceutical  
 compositions, e.g. vaccines, for the diagnosis and treatment of  
 pancreatic cancer. Such compositions may be useful for inhibiting the  
 development of cancer in a patient, or as markers for the progression

CC of cancer. The polynucleotide sequences may also be used as probes  
 CC or primers for nucleic acid hybridisation assays. ABK44061-ABK44209  
 CC represent cDNA sequences encoding for human pancreatic tumour proteins.

XX Sequence 1471 BP; 423 A; 317 C; 353 G; 378 T; 0 other;

Query Match 24.4%; Score 342.6; DB 24; Length 1471;  
 Best Local Similarity 57.0%; Pred. No. 2.6e-86;  
 Matches 791; Conservative 0; Mismatches 569; Indels 27; Gaps 8;

QY 29 TGTCTTTGGCAGCATCAAGAGGAAGAAGTTTCTATGAAGTTAGGTTTCAAG 88  
 |||||  
 Db 38 TGCTGCTGGGAGCAGTAGCAGGAAAAGAAGTTTGTACGAAAGACGCGCTGCTTCAGTG 97  
 |||||  
 QY 89 ATGTTTACCATGACACAGGAGACTTCTCAACAGAGTTGGTAGGTTTACCCCTGGTCTCCAG 148  
 |||||  
 Db 98 ATGACTCCCATGCTCAGGAATTACGGAAAGACCCCTCCATATATTGCTTGGTCTCCAA 157  
 |||||  
 QY 149 AGAGATTAACACTCGTTTCTGCTCTACACTATACAAATCCAAATGCCATTCAGGAGA 208  
 |||||  
 Db 158 AAGATGTCAACACCCCGCTTCTCTCTATATATAATGAGAAACCCCAACAACTTTCAAGAAG 217  
 |||||  
 QY 209 TCAGTGGCGTTAATTCTTCAACTATCCAAGCCTCATATTTTGGAAACAGACAAGATCACCC 268  
 |||||  
 Db 218 T---TGCGCAGATTCTATCAGCNATCAGTGGCTCCATTTCAAAAACAATAGAAAACCTC 274  
 |||||  
 QY 269 GTATCAACATAGCTGGAT-----GGAAAAACAGATGGCAAAATGGCAGAGAGACATGTGCA 322  
 |||||  
 Db 275 GCTTTATTATTCTAGGATTCTATAGACAAGGGAGAGAAGAAACTGGCTGGCCAATGTGTGCA 334  
 |||||  
 QY 323 ATGTGTTGCTACAGCTGGAAGATATAAATTGCAATTAATTAGATTGATCAACGGTTTCAAC 382  
 |||||  
 Db 335 AGAATCTGTTCAAGTGGAAAGTGTGAACGTATCTGTGTGGACTGGAAGGTGGCTTCCC 394  
 |||||  
 QY 383 GG---GAATACATCCATGCTGTAAACAATCTCCGTTGTTGTGTGCTGAGGTGGCTTATT 439  
 |||||  
 Db 395 GAACTGGATACACACAAGCCTCGCAGAAACATCAGGATCGTGGGAGCAGAAAGTGGCATATT 454  
 |||||  
 QY 440 TTATTGATGTTCTCATGAAAAAATTTGAATATTTCCCTTCTAAAGTGCACCTTGATTGGCC 499  
 |||||  
 Db 455 TTGTGAAATTTCTTCAGTCCGCGTTCCGTTTACTCACCTTCCAAACGTGCATGTCAATTGGCC 514  
 |||||  
 QY 500 ACAGCTTGGGAGCACACCTGGCTGGGGAAGCTGGGTCAAGGATACCAGG---CCTTGGAA 556  
 |||||  
 Db 515 ACAGCTTGGGTGCCACCGCTGCTGGGAGGCTGGAAGAGAACCAATGGGACCATTTGGAC 574  
 |||||  
 QY 557 GAATAACTGGGTTGGACCCAGCTGGGCCATTTTTCACAACTCCCAAGAGGAAGTCAGGC 616  
 |||||  
 Db 575 GCATCACAGGGTTGGACCCAGCAGAACCTTGTCTTCAGGGCACACCTGAATTAGTCCGAT 634  
 |||||  
 QY 617 TAGACCCCTCGGATGGCCAACTTTGTTGACGTTTATTCATAAAATGCAGCTCGCATCCTT 676  
 |||||  
 Db 635 TGGACCCCGAGGATGCCAAATTTGTGGATGTAATTCACAGGATGTGCCCCCATAGTCC 694  
 |||||  
 QY 677 TTGAGCTGTTGTTGGAAACCATGATGCTTGTGGTCACTTTGATCTTTTACCCAAATGGAG 736  
 |||||  
 Db 695 CCAATTTGGGTTTGGAAATGAGCCAAAGTCGTGGGCCACCTAGATTTCTTTCCAAATGGAG 754  
 |||||  
 QY 737 GGAAGCATGCCAGGATGTGAAGACTTAATTACACCTTTACTGAAATTTAATTCAATG 796  
 |||||  
 Db 755 GAGTGGAAATGCCTGGATGTAAAGAAAGAACATT---CTCTCAGATTGTGGACATAGACG 811  
 |||||  
 QY 797 CTTACAAAAAAGAAATGGCTTCTCTTTGACTGTAAACCATGCCGGAAGTTATCAATTTT 856  
 |||||  
 Db 812 GAATCTGGGAAGGACCTCGAGACTTTTGGCGSCCTGTAATCACTTAAAGAGCTACAAATATT 871  
 |||||  
 QY 857 ATGCTGAAGCAATCTTAATCTGTAGTCATTATTCTTATCTCTGTAGATCTACACAT 916  
 |||||  
 Db 872 ACACGTAGACATGTCACACCTGATGGCTTTGCTGGATTCCTGCTGCTCTTCAACG 931  
 |||||  
 QY 917 CTTTAAAGCAGGAAATTTGCTTTTGTTCAAAAGAGGTTGCCCAACAATGGGTCAAT 976  
 |||||  
 Db 932 TCCTTCACTGCAACAAAGTGTTCCTTGTCTCAAGTGGAGGCTGCCCCACAGATGGGTCACT 991  
 |||||



977 TTGCTGATAGATTTCACCTTCAAAATATGAAGACTAATGGATCACATTAATTTTAAACA 1036  
992 ATGCTGATAGATATCTCTGGGAAA---ACAAATGATGGGCCAGAAATTTTATCTAGACA 1048  
1037 CAGGCTCCCTTTCCCAATTTGCCGTTGGGACACAAATGCTGTAACTCAGTGGAA 1096  
1049 CTGGTATGCCAGTAATTTTGCAGTTGGAGGTATAAGGTATCTGCACACTGCTGGAA 1108  
1097 GCGAAGTCACTCAAGGAAGTCTTCTTCTGTGTAGCGGGGCAATTTGGAAACTGGGG 1156  
1109 AAAAGGTTAC---AGGACACATACTAGTTCTTGTTCGGAATAAAGGAACCTTAAGC 1165  
1157 AGTTGCCAATGTCAGTGGAAACTTGGCCAGGCATGCTTACACAAATTAATCGATG 1216  
1166 AGTATGAATTTTCAAGGGCACTCTCAACACGATAGTACTCTTCCAATGAATTTGACT 1225  
1217 CAGATGTTTAAAGTTGGAAACATTAACAAGTGTTCAGTTTCATCTGGAATAACATTTG 1276  
1226 CAGATGTGGATGTTGGGACTTGCAGATGGTTAAATTTATTTGGTATAACATGTGATCA 1285  
1277 AAGATTCACAAATAGTTGGGACAGAAATGGTGATAAATACATCTGGGAATATGGAT 1336  
1286 ACCCAACTTTACCTAGATGGGAGACTCCAAGATTAAGTGGAGAC---AAATGTTGGAA 1342  
1337 ATAAATCTACCTTCTGTAGCCAGACATTAATGGACCTAATATTTCTCCAGAACCTGAAC 1396  
1343 AACAGTTCAACTTCTGTAGTCCAGAAACCTCAGGGAGGAGTTCTGCACCTCACAC 1402  
1397 CATGCTA 1403  
1403 CGTGTTA 1409

LT 12

9968

ABL69968 standard; DNA; 1471 BP.

ABL69968;

15-MAY-2002 (first entry)

Pancreas cancer related gene sequence SEQ ID NO:8305.

Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
gene; ds.

Homo sapiens.

WO200194629-A2.

13-DEC-2001.

30-MAY-2001; 2001WO-US10838.

05-JUN-2000; 2000US-209473P.

05-JUN-2000; 2000US-209531P.

18-SEP-2000; 2000US-23133P.

18-SEP-2000; 2000US-231617P.

20-SEP-2000; 2000US-234009P.

20-SEP-2000; 2000US-234034P.

20-SEP-2000; 2000US-234052P.

22-SEP-2000; 2000US-234509P.

22-SEP-2000; 2000US-234567P.

25-SEP-2000; 2000US-234923P.

25-SEP-2000; 2000US-234924P.

25-SEP-2000; 2000US-235077P.

25-SEP-2000; 2000US-235082P.

25-SEP-2000; 2000US-235134P.

25-SEP-2000; 2000US-235280P.

26-SEP-2000; 2000US-235637P.

26-SEP-2000; 2000US-235638P.  
27-SEP-2000; 2000US-235711P.  
27-SEP-2000; 2000US-235720P.  
27-SEP-2000; 2000US-235840P.  
27-SEP-2000; 2000US-235863P.  
28-SEP-2000; 2000US-236028P.  
28-SEP-2000; 2000US-236032P.  
28-SEP-2000; 2000US-236033P.  
28-SEP-2000; 2000US-236034P.  
28-SEP-2000; 2000US-236109P.  
28-SEP-2000; 2000US-236111P.  
28-SEP-2000; 2000US-236842P.  
29-SEP-2000; 2000US-236891P.  
02-OCT-2000; 2000US-237172P.  
02-OCT-2000; 2000US-237173P.  
02-OCT-2000; 2000US-237278P.  
02-OCT-2000; 2000US-237294P.  
02-OCT-2000; 2000US-237295P.  
02-OCT-2000; 2000US-237316P.  
03-OCT-2000; 2000US-237425P.  
03-OCT-2000; 2000US-237598P.  
03-OCT-2000; 2000US-237604P.  
03-OCT-2000; 2000US-237606P.  
03-OCT-2000; 2000US-237608P.  
01-NOV-2000; 2000US-244867P.  
01-NOV-2000; 2000US-245084P.

(AVAL-) AVALON PHARM.

Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

Soppet DR, Weaver Z;

WPI; 2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a  
chemical agent to be tested for anti-neoplastic activity, and  
determining a change in expression of a gene of a signature gene set -

Claim 1; SEQ ID 8305; 44pp; English.

The present invention describes a method (M1) for screening for an  
anti-neoplastic agent. The method involves exposing cells to a chemical  
agent to be tested for anti-neoplastic activity, determining a change in  
expression of at least one gene (I) of a signature gene set, where (I)  
comprises a sequence (S) selected from 8447 sequences (given in AB161664  
to AB170110), or is at least 95% identical to (S), where a change in  
expression is indicative of anti-neoplastic activity. (I) has cytostatic  
activity and can be used in gene therapy. M1 can be used for screening  
an anti-neoplastic agent, and can be used for producing a product which  
is the data collected with respect to the anti-neoplastic agent as a  
result of M1, and the data is sufficient to convey the chemical  
structure and/or properties of the agent. M1 can be used in the  
treatment of cancer such as colon, breast, stomach, lung, thyroid,  
oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
carcinoma, papillary carcinoma and Wilm's tumour.

Sequence 1471 BP; 423 A; 317 C; 353 G; 378 T; 0 other;

Query Match 24.4%; Score 342.6; DB 24; Length 1471;

Best Local Similarity 57.0%; Pred. No. 2.6e-86;

Matches 791; Conservative 0; Mismatches 569; Indels 27; Gaps 8;

Qy 29 TGTTCCTTGGCCACATCAAGAGGAAAGAGTTTGTCTATGAAAGGTTAGGTTGTTCAAAG 88

Db 38 TGCTGCTGGAGCAGTAGCAGGAAGAAGTTTGTACGAAAGACTCGGCTGCTCAGTG 97

Qy 89 ATGCTTACATGGACAGGACTTTCTCAACAGAGTTGGTAGTTTACCTGGTCTCCAG 148

Db 98 ATGACTCCCATGGTTCAGGAATACGAAAGACCCCTCCATATATTCCTGTTGTTCCAA 157

Qy 149 AGAAGATAAACACTCGTTCTCTGCTCTACACTATACACAATCCCAATCCCTATCAGGAGA 208

158 AAGATGTCACACCGCTTCTCCTATATACATTAATGAGAACCCAAACAACTTCAAGAAG 217  
 209 TCAGTGGCGTTAATCTTCAACATATCAAGCCTCATATTTTGGACAGACAGATCAACC 268  
 218 T---TGCGCGAGATTATCAAGCATCAGTGGCTTCAATTTTCAAAACAATAGAAAATC 274  
 269 GTATCAACATAGCTGGAT-----GGAACACAGATGGCAAAATGGCAGAGACATGTGCA 322  
 275 GCTTATATTTATGATTCATAGATCATAGACAAGGAGAGAAAACCTGGTGGCCAATGTGTGA 334  
 323 ATGTGTTGTACAGCTGGAGATATATAATTCATTAATTTAGATTCAGTCAACGTTTCAAC 382  
 335 AGAATCTGTTCAAGGTGGAAAGTGTGAACATGTATCTGTGTGACTGGAAAGTGGCTCCC 394  
 383 GG---GAATACATCTCATGTGTAAACAAATCTCGTGTGTGTGTGCTGAGTGGCTTATT 439  
 395 GAATCGATACACACAAGCTCGCAGACATCAGGATCGTGGGAGCAGAAAGTGGCATATT 454  
 440 TTATTTGATGTTCTCATGAAMAAATTTGAATATTTCCCTTCTTAAAGTGCATTTGATGGCC 499  
 455 TTGTGTAATTTCTTCAGTGGCGCTTCGGTTACTACCTTCCAACTGTCATGTCATTGGCC 514  
 500 ACAGCTTGGGAGCACACCTGGCTGGGGAAGCTGGGTCAAGGATACACAG---CCTTGGAA 556  
 515 ACAGCTGGTGGTCCCGCTGCTGGGAGGCTGGAGGAGAACCAATGGACCATTTGGAC 574  
 557 GAATACTGGGTTGGACCCAGCTGGGCCCATTTTTCACAACACTCAAAGGAAGTCAAGC 616  
 575 GCATCAACAGGTTGGACCCAGCAGAACCTTGTCTTTCAGGGCACAACCTGAATAGTCCGAT 634  
 617 TAGACCCCTCGGATGCCAATTTGTTGAGCTTATTCATACAAATGCAGCTGCCCTCCT 676  
 635 TGACCCCAAGCATGCAATTTTGTGAGTGTATTTCAACGGATGGTGGCCCATATGTC 694  
 677 TTGAGCTGTGTGGAACCAATATGATGTGTGTGCTCATCTTGAATTTTACCAAAATGGAG 736  
 695 CCAATTTGGGTTTGAATGAGCAAGTGGTGGCCACCTAGATTTCTTCCAAATGGAG 754  
 737 GGAAGACATGCCAGATGGAAGCTTAAATACACCTTTACTGAAATTTAACTTCAATG 796  
 755 GAGTGAATGCTCGATGTAAAGAAACATTT---CTCTCTCAGATTTGTGACATAGACG 811  
 797 CTTACAAAAGAAATGGCTTCTCTTTGATGTAAACCAATGCCCGAAGTTATCAATTTT 856  
 812 GAATCTGGGAAGGACTCGAGACTTTTGGCGCTGTAACTACTTAAGAGCTACAATATT 871  
 857 ATGCTGAAGCATTTAATCTGTATGATGATTTATTTGCTTATCTTGTATAGCTTACAT 916  
 872 AACTGATAGCATCGTCAACCCCTGATGGCTTTTGTGGAATTCCTCTGTGCTTACACG 931  
 917 CTTTAAAGCAGGAATTTGCTTCTTTTGTTCAAAAGAGGTTGGCCCAACATGGTCAAT 976  
 932 TCTTCTACTGCAACAGATGTTTCCCTTGTTCGAAGTGGAGGCTGCCACAGATGGGTCACT 991  
 977 TTGCTGATGATTTCACTTCAAAAATATGAAGCAATATGATCAATTTATTTTAAACA 1036  
 992 ATGCTGATAGATATCTCGGAAA---ACAAATGATGTGGCCAGAAAATTTTATCTAGACA 1048  
 1037 CAGGTCCTTCTTCCCATTTGCCGTTGGAGGCAAAATGCTGTTTAACTCAGTGGAA 1096  
 1049 CTGGTGAAGCCAGTAATTTTGCACGTTGGAGGTATGAAGTATCTGTCACTGTCTGGAA 1108  
 1097 GCGAAGTCACTCAAGGAATCTGTTTCTTCTGTAGGGGGGCAATGGGAAAACCTGGG 1156  
 1109 AAAAGGTTAC---AGGACACATAGTATTCTTTTGTTCGGAATAAGAACTTAAGC 1165  
 1157 AGTTTGGCATTTGAGTGAAGAACTTGACCCAGGATGACTTACACAAAATTAATCGATG 1216  
 1166 AGTATGAATTTTCAAGGGCACTCTCAAAACAGATAGTACTCAATCCAATGAATTTGACT 1225  
 1217 CAGATGTTAACTGTGGAACATTTACAAGTGTTCAGTTTCACTCTGGAAGAAAACATTTGTTG 1276

Db 1226 CAGATGCGATGTTGGGACTTGCAGATGGTTAAATTTATTATTGGTATAACAATGTGATCA 1285  
 QY 1277 ACGATTTCTCAGATAGTTGGGAGCAGAAATGGTGAATAAATACATCTCGGAAATATGAT 1336  
 Db 1286 ACCCAACTTTACTAGTGGGAGCATCCAGATTTATAGTGGAGAC---AAATGTTGAA 1342  
 QY 1337 ATAAATCTACTCTTCTGTAGCCAGACATTTATGGGACCTTAATATTCTCCAGAACCTGAAAC 1396  
 Db 1343 ACAGTTTCAACTTCTGTAGTCAGAAACCGTCAGGAGGAGAAATTTCTGCTCACCTTCACAC 1402  
 QY 1397 CATGCTA 1403  
 Db 1403 CGGTGTA 1409

RESULT 13  
 AAC98880  
 ID AAC98880 standard; cDNA; 1536 BP.  
 AC AAC98880;  
 XX  
 DT 09-MAR-2001 (first entry)  
 XX  
 DE Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:108.  
 XX  
 KW Human; pancreas; pancreatic cancer; pancreatic cancer antigen;  
 KW detection; diagnosis; identification; cytostatic; neuroprotective;  
 KW neoplastic; immunomodulatory; relaxant; contraceptive; gynaecological;  
 KW antiinflammatory; cardiac; gene therapy; chromosome mapping;  
 KW linkage analysis; tissue identification; tissue typing; forensic;  
 KW neural; immune system; muscular; reproductive; gastrointestinal;  
 KW pulmonary; cardiovascular; renal; proliferative; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200055320-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 08-MAR-2000; 2000WO-US05989.  
 XX  
 PR 12-MAR-1999; 99US-0124270.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM;  
 XX  
 DR WPI; 2000-579444/54.  
 DR P-PSDB; AAB54115.  
 XX  
 PT New nucleic acid that is a pancreatic cancer antigen for preventing,  
 PT treating, or ameliorating a medical condition, particular pancreatic  
 PT cancer, or for use in assays for diagnosing a pathological condition -  
 XX  
 PS Claim 1; Page 573; 1379pp; English.  
 XX  
 CC AAC98773 to AAC99231 encode the human pancreatic cancer associated  
 CC proteins, called pancreatic cancer antigens, given in AAB54008 to  
 CC AAB54466. The human pancreatic cancer antigens have cytostatic,  
 CC neuroprotective, neoplastic, immunomodulatory, relaxant, contraceptive,  
 CC gynaecological, cardiac and antiinflammatory activities, and can be used  
 CC in gene therapy. The polynucleotide and proteins can be used for  
 CC preventing, treating, or ameliorating a medical condition or in assays  
 CC for diagnosing a pathological condition or a susceptibility to one in a  
 CC subject. Binding partners to the proteins and the activity of the  
 CC proteins can be identified. The pancreatic cancer antigens can be used to  
 CC detect, treat or prevent pancreatic disorders, especially cancer.  
 CC Agonists and antagonists to the antigens can be screened for. The  
 CC pancreatic cancer antigen polynucleotides can be used to design nucleic  
 CC acid hybridisation probes that can be used in chromosome mapping, linkage  
 CC analysis, tissue identification and/or typing and a variety of forensic  
 CC and diagnostic methods. The proteins can be used to generate antibodies  
 CC which are used to purify, detect and target the polypeptides, including

both in vivo and in vitro diagnostic and therapeutic methods. The proteins can be used to treat or prevent neural, immune system, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal or proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent sequences used in the exemplification of the present invention.

Sequence 1536 BP; 476 A; 321 C; 358 G; 380 T; 1 other;

ery Match 24.4%; Score 342.6; DB 21; Length 1536;  
 st Local Similarity 57.0%; Pred. No. 2.7e-86;  
 tches 791; Conservative 0; Mismatches 569; Indels 27; Gaps 8;

29 TGTCTTTGGCAGATCAAGAGAAAGAGTTTCTATGAAGTTAGGTTGTTCAAG 88  
 51 TGCTCTGGAGAGTAGCAGAGAAAGAGTTTCTACGAAGACTCGGCTGCTTCAGTG 110  
 89 ATGTTTACCAATGACAGGACTTTCTCAACAGAGTTGGTAGGTTTACCTGTCTCCAG 148  
 111 ATGACTCCCATGGTCAGGAATTACGGAAGAGCCCTCCCATATATTGGCTTGGTCTCCAA 170  
 149 AGAAGATAAACAACCTGTTCTGCTCTACACTATACAAATCCCAATGCCATCAGGAGA 208  
 171 AAGATGTCAAACCCGCTTCTCTATATATCTATGAGAACCCCAACAACTTTCAAGAAG 230  
 209 TCAGTCCGTTAAATTTCTCAACTATCCAGGCTCATATTTTGAACAGACAGATCAACC 268  
 231 T---TGCGCGAGATTCATCAGATCAGTGGCTCCAATTTCAAAACAAATGAAAACCTC 287  
 289 GTATCAACATAGCTGGAT-----GGAACAGATGGCAATGGCAGAGACATGTGCA 322  
 288 GCTTTATTATTCAATGGATTCTATAGCAAGGAGAGAAAACCTGGCTGGCCAAATGTGTGCA 347  
 323 ATGTGTTGCTACAGCTGGAAGATATAATTCATTAATTTAGATTGSAATCAACGGTTTCA 382  
 348 AGAATCTGTTCAAGTGGAGAGTGTGAATCTGTCTGTGGACTGGAAGGTGGCTCC 407  
 383 GG---GAATACATCCATCTGTAAACAATCCGTTGTGTGTGTGCTGAGGTGGCTTATT 439  
 408 GAACTGGATACACAAAGCTCGCAGAACATCAGGATCGTGGGAGCAGAGTGGCATATT 467  
 440 TTATTGATGTTCTCATGAAGAAATTTGAATTTCCCTTTCTAAGTGCACCTGATTTGGCC 499  
 468 TTGTTGAATTTCTTTCAGTGGCGTTTCGGTTACTACCTTCCAAAGTGCATGATGGCC 527  
 500 ACAGCTTGGGAGCACACTGGCTGGGAGGCTGGTCAAGATACACAGG---CCTTGGAA 556  
 528 ACAGCTTGGTGGCCACGCTGCTGGGAGGCTGGAAGAGAACCAATGGACCATTTGGAC 587  
 557 GAATPACTGGGTTGGACCCAGCTGGCCATTTTTCACAACTCCAAAGGAGTCAAGC 616  
 588 GCATCACAGGTTGGACCCAGCAGAACCTTCTTCAGGGCACACCTGAAATTAGTCCGAT 647  
 617 TAGACCCCTCGGATGCCAACTTTGTCAGCTTATTCATACAAATGCACTGCGATCCTCT 676  
 648 TGGACCCAGGATGCCAAATTTGTGATGTAATTCACAGGATGTGCGCCCATGATGCC 707  
 677 TTGAGCTTGGTGGAAACCAATGAGCTTGTGTGATCTTTTACCTTTTACCCAAATGGAG 736  
 708 CCAATTTGGGTTTGGAAATGAGCCAGTCTGGGCCACCTAGATTCTTTCCAAATGGAG 767  
 737 GGAACACATGCCAGGATGTGAAGACTTAATTACACTTTTACTGAAATTAACTTCAATG 796  
 768 GAGTGGAAATGCCCTGGATGTAAAGAACATTT---CTCTCAGATTGTGGACATAGAGC 824  
 797 CTTACAAAAGAAATGGCTTCTTCTTTGACTGTAAACCATGCCGGAAGTTATCAATTTT 856  
 825 GAATCTGGGAAGGACTCGAGACTTTGGCGCTGTATCACTTAAGAGCTACAATATT 884  
 857 ATGCTGAAGCAATTTTAATCTGATGCAATTTATGCTTATCTTGTAGATCTACACAT 916  
 885 ACATGTATAGATCGTCAACCCCTGATGGCTTGTGGAATTTCCCTGTGCTCTTACAGC 944  
 917 CTTTAAAGCAGGAAATGCTTCTTTTGTTCAAAGAGAGTTGCCCAACAATGGGTCAATT 976

Db 945 TCITCACTGCAAAACAGTGTTCCTTGTCCAAAGTGGAGGCTGCCACAGATGGTCACT 1004  
 Qy 977 TTGCTGATGATTTCACTTCAAAATATGAAGACTAATGATCACTATTTTAAACA 1036  
 Db 1005 ATGCTGATGATATCTTGGGAAA---ACAAATGATGGGCGGCAATTTTATCTAGACA 1061  
 Qy 1037 CAGGTCCTTTCCTCCATTTGCCCCGTTGGAGGCACAAATTTGCTTAAACTCAGTGGAA 1096  
 Db 1062 CTGCTGATGCCAGTAATTTTGACGTTGGAGGTATAAGGTATCTGTACACTGTCTGGAA 1121  
 Qy 1097 GCGAGTCACTCAAGAACTGCTTTCTTCTGTAGGCGGCAATTTGGAAAACCTGGGG 1156  
 Db 1122 AAAAGGTTAC---AGGACACATACTAGTTTCTTGTGGAATAAAGGAAACTCTAAGC 1178  
 Qy 1157 AGTTTGCATTTGTCAGTGGAAACTTTGAGCCAGGATGATCTTACACAAATTTAAATCGATG 1216  
 Db 1179 AGTATGAAATTTTCAAGGCACTCTCAACAGAGATAGTACTCATTCATGATTTGACT 1238  
 Qy 1217 CAGATGTTACGTTGGAACATTAACAAGTGTTCAGTTTCACTGGAACAAACATTTGTTG 1276  
 Db 1239 CAGATGTTGATGTTGGGACTTTCAGATGTTTAAATTTATTTGGTATAACATGTCATCA 1298  
 Qy 1277 AAGATCTCAGATAAAGTTGGGAGCAGAAATGCTGATAAATACATCTGGGAATATGGAT 1336  
 Db 1299 ACCCACTTTACCTAGTGGGAGCATCCAGATTTATAGTGGAGAC---AAATGTTGAA 1355  
 Qy 1337 ATAAATCACTTCTGTAGCCAGACATTTATGGACCTTAATTTCTCCAGAACCTGAAAC 1396  
 Db 1356 AACAGTTCAACTTCTGTAGTCAGAAACCGTCAGGAGGAAAGTTCTGCTCACCTCACAC 1415  
 Qy 1397 CATGCTA 1403  
 Db 1416 CGTGTTA 1422

RESULT 14  
 AAS81083  
 ID AAS81083 standard; cDNA; 1410 BP.  
 XX AAS81083;  
 XX 13-FEB-2002 (first entry)  
 XX DNA encoding novel human diagnostic protein #16887.  
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX Homo sapiens.  
 XX WO200175067-A2.  
 XX 11-OCT-2001.  
 PF 30-MAR-2001; 2001WO-US08631.  
 XX 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX (HYSE-) HYSEQ INC.  
 XX Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 DR P-PSDB; ABG16896.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity  
 XX Claim 1; SEQ ID No 16887; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridization probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).

Sequence 1410 BP; 384 A; 320 C; 369 G; 337 T; 0 other;

Very Match 24.1%; Score 338; DB 23; Length 1410;  
 1st Local Similarity 54.9%; Pred. No. 5.2e-85;  
 Matches 778; Conservative 0; Mismatches 620; Indels 18; Gaps 5;

```

1 ATGCTGGAAATTGGATTGTTCATCTTCTTTGGCACATCAAGAGGAAAGAAGTT 60
1 ATGCTGCCCTTGGACCTCGGCTTCTCTGCTGGCCACAGTCAGAGGAAAGAGTTC 60
61 TCGTATGAAGGTTAGGCTGTTCAAGATGGTTTACATGGACACAGACTTTCACA 120
61 TCGTACGACAACTTGGCTGCTTTCTGATGAAACCATGGCGAGGACCTTCAGCG 120
121 GAGTGTGGAGTTTACCTGGTCTCCAGAGAGATAAACAATCGTTTCTGCTCACT 180
121 CCTGTAAATTAATCTTCCCTGGTCCCGGAGACATTCACCCGCTTCTCTGTAC 180
181 ATACACATCCCAATGCCATCAGAGATCAGTGGGTAAATCTTCAACTATCCAGCC 240
181 AATGAAATCCAAACAATCTCAACTAATCACTGGCACGGAACACAGACCAATGAG 240
241 TCATATTTTGAACAGACAGATCACCCTATCAACATAGCTG-----ATGGA 294
241 TCAACTTCCACTGACCGCAGACAGCTTCATCATCCATGCTTTTAGACAGGCG 300
295 GATGCAATGGCAGAGACAGATGCAATGTGTTGCTACAGCTGGAAGATATAAATTG 354
301 GAGGACAGCTGGCCATCGACATGTGCAAGAAATGTTTGAAGTGGAGAGGTGA 360
355 ATTAATTTAGATTGGATCAACGGTTCACGGAA---TACATCCATGCTGTAACAT 411
361 ATCTGTGGACTGAGGCA CGGTTCGGGCAATGTACACCAAGCCGTGCAAAACAT 420
412 COTGTGTGTGTGCTGAGGTGGCTTATTTTATGATGTTCTCATGAAATAATTTGA 471
421 CGGGTGTGTGGGCGGAGACAGCTTTCTTAATACAGCACTGTGCGAGCTAGGTT 480
472 TCCCTTCTTAAGTGCATTTAATTTGGCCACAGCTTGGAGCACACCTGGCTGG 531
481 AGCTTTGAGGACGTGTCATGTATCGGCGACAGCTGGGCGCGCACACGCGCG 540
532 GGGTCAAGGATACAGGCG---TTGGAAGATAACTGGGTGGACCCAGCTGGCCAT 588
541 GGCAGAGGCTGGGGGCGGCTGGGCGAGATACAGGGCTGGATCCAGAGGCGCG 600
589 TTCACACACTCCAAAGGAGTCAAGCTAGACCCCTCGGATGCCAACTTTGTGAG 648
601 TTCAGGATGAACCTGAGGAGTTTCGGTTGGATCCATCTGACCGCTGTTTGTG 660
  
```

```

QY 649 ATTACATAAATGACAGTCGCATCTCTTTGAGCTTGGTGTGGAAACCATGTGCTGT 705
Db 661 ATTACACAGATTCCTTCCCATAGTCTCTTCCTAGTTCGGAAATGAGCCAAAGGT 720
QY 709 GGTCACTGTGACTTTTACCACAAATGGAGGAGACACATGCCAGGATGTGAAGACTTA 768
Db 721 GGCCATCTGGATTTCTTCCAAATGGAGAAAGGAATGCCCGATGTAGAA---AA 777
QY 769 ACACCTTTACTGAAATTTAACTCAATGCTTCAAAAAGAAATGGCTTCTCTTTGAC 828
Db 778 GTCCCTTCAACATTAATGATGATGGAATATGGAAGAAATGTGGCTTTGTGTCT 837
QY 829 TGTAAACATGCCGAGTTATCAATTTTATGCTCAAAAGCATTTCTTAATCTGTGCA 888
Db 838 TGAATCACTCAAGAGCTTCGAGTATTACTCAAGCAGCGTCTCAACCTGATGGCTTC 897
QY 889 ATTGCTTATCTTGTAGATCTCAACATCTTTTAAAGCAGAAATTCCTTCTTTGTTC 948
Db 898 CTGGCTATCCCTGTGCTCTCTAGATGAGTTTCAGGAGAGTAAGTGTTCCTTGTCCA 957
QY 949 AAAGAAGGTTGCCCAACAATGGGTCAATTTGCTGATAGATTTCACCTTCAAAAATATGA 1008
Db 958 GCTGAAGGATGCCCAAAATGGGCACTATGCTGACCAATTTAAGGGGAAA---ACA 1014
QY 1009 ACTAATGATCACATTAATTTTAAACACAGGGTCCCTTTCCCATTTGGCCCTTGGAG 1068
Db 1015 GCTGTGGAACAAACCTTTTCTGTAACACAGGAGAGAGTGGTAACCTTACTAGTTG 1074
QY 1069 CACAAATGCTGTCTTAAACTCAGTGGAGGAGCACTCACTCAAGGAACTGTCTTCTCG 1128
Db 1075 TATAGGTATCAGTCACACTTCTGGAAGAGAAAGTGAATGGGTACATCAGGATG 1134
QY 1129 GTAGCGGGGCAATTTGGGAAACCTGGGAGCTTTGCCATTTCTAGTGGAAACCTTGAG 1188
Db 1135 TTGTATGGAAGTAAATGAAACTCGAAACAATATGAGATTTTCAAGGATCCCTCAAC 1194
QY 1189 GGCATGCTTACACAAATTAATCGATGACAGATGTTAACTGGAACATTAACAAGTGT 1248
Db 1195 GATCAAGTACACGCTGTGTATGATGTGGATTTTAAATGTGGAAAAATACAGAAG 1254
QY 1249 CAGTTCATCTGGAAGAAACATTTGTTTGAAGATTTCTCAGAAATGTTGGGAGCAGA 1308
Db 1255 AAATTCCTCTGGAACAAACCTGGGATATAATCTATCTGACCCCAACTGGGGCTT 1314
QY 1309 GTGATAATACATCTGGGAAATATGGATATAAATCTACCTTCTGTAGCCAAAGCAT 1368
Db 1315 ATCAGGTGCAAGTGGTGAAGATGGGACTGAGTATAATTTTGTAGCAGCAGACACT 1374
QY 1369 GGACCTAATATTCTCCAGAACCTGAAACCATGCTAA 1404
Db 1375 GAAGAAACGCTCTTGCAATCTCTTTACCCCTTGTAA 1410
  
```

# RESULT 15

ABL69986  
 ID ABL69986 standard; DNA; 1450 BP.  
 XX ABL69986;  
 AC ABL69986;  
 XX  
 DT 15-MAY-2002 (first entry)  
 XX  
 DE Pancreas cancer related gene sequence SEQ ID NO:8323.  
 XX  
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
 KW cytosolic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;  
 XX gene; ds.  
 OS Homo sapiens.  
 XX  
 FN WO200194629-A2.  
 XX  
 PD 13-DEC-2001.

30-MAY-2001; 2001WO-US10838.

05-JUN-2000; 2000US-209473P.  
05-JUN-2000; 2000US-209531P.  
18-SEP-2000; 2000US-233133P.  
18-SEP-2000; 2000US-233617P.  
20-SEP-2000; 2000US-234009P.  
20-SEP-2000; 2000US-234034P.  
20-SEP-2000; 2000US-234052P.  
22-SEP-2000; 2000US-234059P.  
22-SEP-2000; 2000US-234567P.  
25-SEP-2000; 2000US-234923P.  
25-SEP-2000; 2000US-234924P.  
25-SEP-2000; 2000US-235077P.  
25-SEP-2000; 2000US-235082P.  
25-SEP-2000; 2000US-235134P.  
25-SEP-2000; 2000US-235280P.  
26-SEP-2000; 2000US-235637P.  
26-SEP-2000; 2000US-235638P.  
27-SEP-2000; 2000US-235711P.  
27-SEP-2000; 2000US-235720P.  
27-SEP-2000; 2000US-235840P.  
27-SEP-2000; 2000US-235863P.  
28-SEP-2000; 2000US-236028P.  
28-SEP-2000; 2000US-236032P.  
28-SEP-2000; 2000US-236033P.  
28-SEP-2000; 2000US-236034P.  
28-SEP-2000; 2000US-236109P.  
28-SEP-2000; 2000US-236111P.  
29-SEP-2000; 2000US-236844P.  
29-SEP-2000; 2000US-236891P.  
02-OCT-2000; 2000US-237172P.  
02-OCT-2000; 2000US-237173P.  
02-OCT-2000; 2000US-237278P.  
02-OCT-2000; 2000US-237294P.  
02-OCT-2000; 2000US-237295P.  
02-OCT-2000; 2000US-237318P.  
03-OCT-2000; 2000US-237423P.  
03-OCT-2000; 2000US-237598P.  
03-OCT-2000; 2000US-237604P.  
03-OCT-2000; 2000US-237606P.  
03-OCT-2000; 2000US-237609P.  
01-NOV-2000; 2000US-244867P.  
01-NOV-2000; 2000US-245084P.

(AVAL-) AVALON PHARM.

Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
Soppet DR, Weaver Z;

WPI; 2002-189264/24.

Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set -

Claim 1; SEQ ID 8323; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 847 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer,

CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilm's tumour.

SQ Sequence 1450 BP; 396 A; 326 C; 379 G; 349 T; 0 other;

Query Match 24.1%; Score 338; DB 24; Length 1450;  
Best Local Similarity 54.9%; Pred. No. 5.3e-85;  
Matches 778; Conservative 0; Mismatches 620; Indels 18; Gaps 5;

QY 1 ATGCTTGGAAATTGGATTGTTGCATCTTCTTTGGGCACATCAAGAGGAAAGAGTT 60  
DB 1 ATGCTGCCCCCTTGGACCTTCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60  
QY 61 TGCTATGAAGGTTAGGGTGTTCAGAGATGGTTTACCATGGACACAGGACTTCTCAACA 120  
DB 61 TGCTACGGACAACTTGGCTGCTTTCTGATGAAAAACCATGGCAGGAAACCTTCAGCGA 120  
QY 121 GAGTTGGTAGGTTTACCCCTGGTCTCCAGAGAGATAAACAACCTCGTTTCTCTCTCTACT 180  
DB 121 CCTGTAAATTTACTTCCCTGGTCCCCCGAGGACATTCACACCGCTTCTCTCTCTACA 180  
QY 181 ATACAAATCCCAATGCTATCAGAGATCAGTGGGTAAATTTCTCACTATCOAGCC 240  
DB 181 AATGAAATCCAAACAACTTCCAACTAATCCTGGCAGGAAACAGACACCACTGAGCT 240  
QY 241 TCATATTTGGAAACAGACAAAGATCACCGGTATCAATAGCTGG-----ATGGAACA 294  
DB 241 TCAACTTCCAACTGGACCCAGACAGCGTTTCATCATCTGCTTCTTAGACAAGGG 300  
QY 295 GATGGCAATGGCAGAGAGACATGTGCAATGTGTTCTACAGCTGGAGATATAAATTGC 354  
DB 301 GAGGACAGCTGGCCATCGGACATGTGCAAGAAATTTTGAAGTGGAGAGGTGAACATGC 360  
QY 365 ATTAATTTAGATTGGATCAACCGTTCAAGGAA---TACATCCATGCTCTAAACAATCTC 411  
DB 361 ATCTGTGTGAGTGGAGGACCGGTCCCGGCAATGTACACCCAGCGTGCACAAACAT 420  
QY 412 CGTGTGTGTGCTGAGGTGGCTTATTTTATGATGTTCTCATGAAAAATTTGAATAT 471  
DB 421 CGGTTGTGGGGCGGAGACAGCTTTCTTAATACAGACACTGTCCAGCAGCTAGGGTAC 480  
QY 472 TCCCCCTTCTAAAGTGCACTTGATTGGCCACAGCTTGGGAGCACACCTGGCTGGGAAGCT 531  
DB 481 AGCTTGGAGACGTGCATGTCATCGGCCACAGCTGGCGCGCACACGCGCGAGGCG 540  
QY 532 GSGTCAAGGATACACAGGCC---TTGGAAGAAATTAATCTGGTTGGACCCAGCTGGCCATT 588  
DB 541 GGCAGGAGCTGGGGGCGCGCGTGGGCGAGATCACAGGCTGGATCCAGCAGGCGCGTGC 600  
QY 589 TTCCACACACTCCAAAGGAAGTCAGGCTAGACCCCTCGGATGCCAACTTTTGTGACGTT 648  
DB 601 TTCCAGGATGAACCTGAGGAGGTTGCGTTGGATCCATCTGACCGCTGTTGTGATGTG 660  
QY 649 ATTATATAATGAGCTCCATCTCTTTGAGCTGGTGTGGAAACCAATGATGCTTGT 708  
DB 661 ATTCACAGAGATCTCTCTCCATAGTTCCTTCCTAGGTTTCGGAATGAGCCAAAGGTG 720  
QY 709 GGTCTATCTTGACTTTTACCAAAATGGAGGAAGCAATCCAGGATGTGAAGACTTAAT 768  
DB 721 GGCCATCTGGATTTCTTCCAAATGGAGGAAGAAATGCCCGGATGTAAGAA---AAAT 777  
QY 769 ACACCTTTACTGAAATTTAATCTCAATGCTTACAAAAGAAATGCTTCTCTTTGAC 828  
DB 778 GTCTCTTCAACCAATTTACTGATATGATGGAATGGAAGAAATGGTGGCTTTGTGCT 837  
QY 829 TGTAAACCATGCGCAAGTATCAATTTATGCTGAAAGCAATTTCTTAATCCTGATGCA 888  
DB 838 TGCATCACCTAAGAACTTCGAGTATTACTCAAGCAGCGTCTCTCAACCTGATGCTTC 897  
QY 889 ATTGCTATCTCTGTAGATCTTACACATCTTTTAAAGCAGGAAATGCTTCTTTGTTCC 948  
DB 898 CTGGGCTATCCCTGTGCT 957

949 AAAGAAGTTGCCCAACAATGGGTCTATTTGCTGATAGATTTCACCTTCAAAAATATGAAG 1008  
958 GCTGAAGGATGCCCAAAATGGGCACTATGCTGACCAATTTAAGGGAAA---ACAAGT 1014  
1009 ACTAATGGATCACATTATTTTAAACACACAGGGTCCCTTTCCCCATTTGCCCGTTGGAGG 1068  
1015 GCTGTGGAAACAACCTTTTCTGAAACACAGGAGAGAGTGGTAACCTTTACTAGTTGGAGA 1074  
1069 CACAAAATGTCTGTAAACTCAGTGGGAAGCGAAAGTCACTCAAGGAACTGTCTTCTTCGT 1128  
1075 TATAAGGTATCAGTCACACTTTCTGGAAAAGAGAAAGTGAATGGGTACATCAGGATTGCT 1134  
1129 GTAGCGGGCAATGGGAAACTGGGGAGTTTGGCCATTGTCTAGTGGAAAACTTGAGCCA 1188  
1135 TTGTATGGAAGTAATGAAACTCGAAACAAATATAGAGATTTTCAAGGATCCCTCAAAACA 1194  
1189 GGCATGACTTACACAAAAATTAATCGATGCAGATGTTAAAGTTGGAACACATTACAAGTGT 1248  
1195 GATGCAAGTCACACGTGTGCTATTGATGTGGATTTTAAGTTGGAAAAATACAGAAAGTT 1254  
1249 CAGTTCATCTGGAAAAAACATTTGTTTGAAGATTCACAGAAATAGTTGGAGCAGAAATG 1308  
1255 AAATTCCTCTGGAACAAACCGTGGGATAAATCTATCTGAGCCCAAACTGGGGCTTCCCAA 1314  
1309 GTGATAAATACATCTGGGAATATGGATATATAATCTACCTTCTGTAGCCCAAGACATTATG 1368  
1315 ATCAGTGCMAAGTGTGAAGTGGGACTGAGTATAATTTTGTAGCAGCAGACACTGTG 1374  
1369 GGACCTAATATCTCCAGAACCTGAAACCATGCTAA 1404  
1375 GAAGAAAAAGCTTTGCAATCTCTTTACCCCTTGTTAA 1410

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

cleic - nucleic search, using sw model

on: January 6, 2004, 08:43:32 ; Search time 480 Seconds  
(without alignments)  
10124.232 Million cell updates/sec

a: US-10-038-517-1

act score: 1404

nce: 1 atccttggatttgattgt.....agaaactgaaccatgctaa 1404

ng table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

hed: 2263443 seqs, 1730637950 residues

number of hits satisfying chosen parameters: 4526886

um DB seq length: 0

um DB seq length: 2000000000

rocessing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

ase : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US08A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US08B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

t	Score	Query Match	Length	ID	Description
1	1404	100.0	1404	14	US-10-038-517-1
2	1402.4	99.9	2352	13	US-10-403-745-1
3	1399.4	99.7	1401	13	US-10-403-745-2
4	1234.6	87.9	1393	13	US-10-312-088-1
5	386.2	27.5	1481	10	US-09-969-347-193
6	342.6	24.4	1471	9	US-09-923-779-145
7	342.6	24.4	1471	10	US-09-969-347-176
8	342.6	24.4	1471	13	US-10-191-987-117
9	342.6	24.4	1536	9	US-09-925-297-108
0	338	24.1	1450	10	US-09-969-347-194
1	338	24.1	1506	9	US-09-925-297-280
2	224	16.0	938	9	US-09-923-779-139
3	209.2	14.9	907	9	US-09-923-779-76
4	204.2	14.5	879	9	US-09-923-779-94
5	203	14.5	963	9	US-09-923-779-22

16	202.6	14.4	766	9	US-09-923-779-73	Sequence 73, Appl
17	198.8	14.2	872	9	US-09-923-779-104	Sequence 104, App
18	194.4	13.8	800	15	US-10-060-036-494	Sequence 494, App
19	192.6	13.7	789	15	US-10-060-036-459	Sequence 459, App
20	189.2	13.5	735	9	US-09-923-779-82	Sequence 82, Appl
21	188.6	13.4	633	9	US-09-925-297-22	Sequence 22, Appl
22	187	13.3	750	9	US-09-923-779-77	Sequence 77, Appl
23	186.2	13.3	764	9	US-09-923-779-78	Sequence 78, Appl
24	179.8	12.8	548	15	US-10-060-036-3160	Sequence 3160, Ap
25	179.8	12.8	567	15	US-10-060-036-3373	Sequence 3373, Ap
26	179.6	12.8	789	9	US-09-923-779-75	Sequence 75, Appl
27	179.4	12.8	599	15	US-10-060-036-3414	Sequence 3414, Ap
28	178.8	12.7	526	15	US-10-060-036-3225	Sequence 3225, Ap
29	178.8	12.7	780	9	US-09-923-779-85	Sequence 85, Appl
30	177.6	12.6	741	9	US-09-923-779-80	Sequence 80, Appl
31	175	12.5	761	9	US-09-923-779-74	Sequence 74, Appl
32	172.2	12.3	659	15	US-10-060-036-4353	Sequence 4353, Ap
33	172.2	12.3	55355	9	US-09-735-933-3	Sequence 3, Appli
34	172	12.3	524	15	US-10-060-036-3545	Sequence 3545, Ap
35	170.8	12.2	752	9	US-09-923-779-87	Sequence 87, Appl
36	170.6	12.2	649	15	US-10-060-036-3573	Sequence 3573, Ap
37	169.4	12.1	804	9	US-09-923-779-88	Sequence 88, Appl
38	168.8	12.0	748	9	US-09-923-779-96	Sequence 96, Appl
39	168	12.0	712	9	US-09-923-779-103	Sequence 103, App
40	166.2	11.8	743	9	US-09-923-779-101	Sequence 101, App
41	163	11.6	528	15	US-10-060-036-4137	Sequence 4137, Ap
42	161.4	11.5	528	15	US-10-060-036-3387	Sequence 3387, Ap
43	161.2	11.5	543	15	US-10-060-036-3564	Sequence 3564, Ap
44	160.2	11.4	511	15	US-10-060-036-3568	Sequence 3568, Ap
45	160.2	11.4	555	15	US-10-060-036-4187	Sequence 4187, Ap

## ALIGNMENTS

### RESULT 1

US-10-038-517-1  
; Sequence 1, Application US/10038517  
; Publication No. US20020115844A1  
; GENERAL INFORMATION:  
; APPLICANT: Yu, Xuanchuan  
; APPLICANT: Turner, C. Alexander Jr.  
; TITLE OF INVENTION: No. US20020115844A1 Human Lipase and Polynucleotides Encoding ti  
; FILE REFERENCE: LEX-0293-USA  
; CURRENT APPLICATION NUMBER: US/10/038,517  
; PRIOR FILING DATE: 2002-01-03  
; PRIOR APPLICATION NUMBER: US 60/259,830  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; TYPE: DNA  
; ORGANISM: homo sapiens  
; US-10-038-517-1

Query Match	100.0%	Score 1404;	DB 14;	Length 1404;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1404;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGCTTGAATTGGATTGTCATTCCTTTGGCAGATCAAGAGGAAAGAGTT	60	
Db	1	ATGCTTGAATTGGATTGTCATTCCTTTGGCAGATCAAGAGGAAAGAGTT	60	
QY	61	TGCTATGAAGTTAGGCTTTCAGAGATGTTTACATGACGACGACTTCTCAACA	120	
Db	61	TGCTATGAAGTTAGGCTTTCAGAGATGTTTACATGACGACGACTTCTCAACA	120	
QY	121	GAGTTGGTAGGTTTACCTTGTCTCCAGAGAGATAAACACTCTCTTCTCTACACT	180	
Db	121	GAGTTGGTAGGTTTACCTTGTCTCCAGAGAGATAAACACTCTCTTCTCTACACT	180	
QY	181	ATACACATCCCATGCTTATCAGAGATCAGTGGGTAAATTTCTTCACTATCCAGGCC	240	



181 |ATACAAATCCAAATGCCATACAGAGATCAGTCGGTTAAATCTTCAACTATCCAGCC 240  
241 |TCATATTTTGGACAGACAGATCACCGGTATCAACATAGCTGGATGGAACAGATGCG 300  
241 |TCATATTTTGGACAGACAGATCACCGGTATCAACATAGCTGGATGGAACAGATGCG 300  
301 |AAATGGCAGAGACATGTGCAATGTGTGCTCAGCTGGAAGATATAAATTCATTAAAT 360  
301 |AAATGGCAGAGACATGTGCAATGTGTGCTCAGCTGGAAGATATAAATTCATTAAAT 360  
361 |TTAGATTGGATCAACGGTTACCGGAATACATCATCTCTAAACATCTCCGTGTGTT 420  
361 |TTAGATTGGATCAACGGTTACCGGAATACATCATCTCTAAACATCTCCGTGTGTT 420  
421 |GGTGTGAGGTGGCTTAATTTATTTGATGTTCTCATGAAAAAATTTGAATATTCCTTCT 480  
421 |GGTGTGAGGTGGCTTAATTTATTTGATGTTCTCATGAAAAAATTTGAATATTCCTTCT 480  
481 |AAATGTCATTTGATTTGGCCAGCTTGGGAGCACCTGCTGGGAGAGCTGGGTCAAG 540  
481 |AAATGTCATTTGATTTGGCCAGCTTGGGAGCACCTGCTGGGAGAGCTGGGTCAAG 540  
541 |ATACAGGCTTGAAGAATAACTGGTTGGACCCAGCTGGGCCATTTTCCACACACT 600  
541 |ATACAGGCTTGAAGAATAACTGGTTGGACCCAGCTGGGCCATTTTCCACACACT 600  
601 |CCAAAGGAAGTCAGCTAGACCCCTCGATGGCAACTTTGTTGACGTTATTCATCAAAAT 660  
601 |CCAAAGGAAGTCAGCTAGACCCCTCGATGGCAACTTTGTTGACGTTATTCATCAAAAT 660  
661 |CGAGCTGGCATCTCTTTGAGCTTGGTTGGAAACCAATGATCTTGTGCTCATCTTGAC 720  
661 |CGAGCTGGCATCTCTTTGAGCTTGGTTGGAAACCAATGATCTTGTGCTCATCTTGAC 720  
721 |TTTTTACCAAAATGAGGAGAGACATGCGAGGATGGAAGCTTAATTTACACTTTACTG 780  
721 |TTTTTACCAAAATGAGGAGAGACATGCGAGGATGGAAGCTTAATTTACACTTTACTG 780  
781 |AAATTTAACTTCAATGCTTACAAAAGAAATGCTTCTTTGACGTGAACCATGCC 840  
781 |AAATTTAACTTCAATGCTTACAAAAGAAATGCTTCTTTGACGTGAACCATGCC 840  
841 |CGAAGTTATCAATTTATGCTGAAAGCAATCTTAATCTTGATGCAATTTATGCTTATCT 900  
841 |CGAAGTTATCAATTTATGCTGAAAGCAATCTTAATCTTGATGCAATTTATGCTTATCT 900  
901 |TGTAGATCTTACATCTTTTAAAGCAGGAAATTTGCTTTTCTTCTTCCAAAGAGTTGC 960  
901 |TGTAGATCTTACATCTTTTAAAGCAGGAAATTTGCTTTTCTTCTTCCAAAGAGTTGC 960  
961 |CCAAATATGGGTCAATTTGCTGATAGATTTCACTTCAAAAATATGAAGACTAATGGATCA 1020  
961 |CCAAATATGGGTCAATTTGCTGATAGATTTCACTTCAAAAATATGAAGACTAATGGATCA 1020  
1021 |CATTAATTTTAAACAGAGGTCCTTTCCCAATTTGCCGTTGGAGGCAAAATTTGCT 1080  
1021 |CATTAATTTTAAACAGAGGTCCTTTCCCAATTTGCCGTTGGAGGCAAAATTTGCT 1080  
1081 |GTTAAACTCAGTGGAGGAGTCACTCAAGGAACCTGCTTTCTTCTGTTAGGGGGGCA 1140  
1081 |GTTAAACTCAGTGGAGGAGTCACTCAAGGAACCTGCTTTCTTCTGTTAGGGGGGCA 1140  
1141 |ATTGGGAAAACCTGGGAGTTTGCATTTGTCAGTGGAAAACTTTGAGCCAGGATGACTTAC 1200  
1141 |ATTGGGAAAACCTGGGAGTTTGCATTTGTCAGTGGAAAACTTTGAGCCAGGATGACTTAC 1200  
1201 |ACAAATTAATTCAGATGAGATTTAAAGTTGGAACATTAACAGTGTTCAGTTCACTCG 1260  
1201 |ACAAATTAATTCAGATGAGATTTAAAGTTGGAACATTAACAGTGTTCAGTTCACTCG 1260  
1261 |AAAAACATTTGTTGAAGATCTCAGAAATAAGTTGGGAGCAGAAATGGTGATAAATACA 1320

Db 1261 |AAAAACATTTGTTGAAGATCTCAGAAATAAGTTGGGAGCAGAAATGGTGATAAATACA 1320  
Qy 1321 |TCCTGGAAATATCGATATAAATCTACCTTCTGTAGCCAGACATTTATGGACCTAATATT 1380  
Db 1321 |TCCTGGAAATATCGATATAAATCTACCTTCTGTAGCCAGACATTTATGGACCTAATATT 1380  
Qy 1381 |CTCCAGAACCTGAAACCATGCTAA 1404  
Db 1381 |CTCCAGAACCTGAAACCATGCTAA 1404

## RESULT 2

US-10-403-745-1  
; Sequence 1, Application US/10403745  
; Publication No. US20030165975A1  
; GENERAL INFORMATION:  
; APPLICANT: Khodadoust, Mehran  
; APPLICANT: Kapeller-Liebermann, Rosana  
; TITLE OF INVENTION: No. US20030165975A1 Human Lipase Proteins, Nucleic Acides Encod  
; FILE OF INVENTION: Them, and Uses of Both of These  
; FILE REFERENCE: 10147-14  
; CURRENT APPLICATION NUMBER: US/10/403,745  
; PRIOR FILING DATE: 2003-03-31  
; PRIOR APPLICATION NUMBER: US/09/411,132A  
; PRIOR FILING DATE: 2000-09-12  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 2352  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (2159)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (2307)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (2313)  
US-10-403-745-1

Query Match 99.9%; Score 1402.4; DB 13; Length 2352;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1403; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 |ATGCTTGGAAATTTGGATTGTTGTCATCTTGTCTTTTGGCACAATCAAGAGGAAAAAGATT 60  
Db 125 |ATGCTTGGAAATTTGGATTGTTGTCATCTTGTCTTTTGGCACAATCAAGAGGAAAAAGATT 184  
Qy 61 |TGCTATGAAAGGTTAGGGTGTTCAAAGATGTTTACCATGACCCAGGACTTTCTCAACA 120  
Db 185 |TGCTATGAAAGGTTAGGGTGTTCAAAGATGTTTACCATGACCCAGGACTTTCTCAACA 244  
Qy 121 |GAGTTGGTGGTTTACCCTGGTCTCCAGAGAGATAAACAATCGTTTCTGCTCTACACT 180  
Db 245 |GAGTTGGTGGTTTACCCTGGTCTCCAGAGAGATAAACAATCGTTTCTGCTCTACACT 304  
Qy 181 |ATACACAATCCCAATGCCCTATCAGAGATCAGTGGGTAAATCTTCAACTATCCAGGCC 240  
Db 305 |ATACACAATCCCAATGCCCTATCAGAGATCAGTGGGTAAATCTTCAACTATCCAGGCC 364  
Qy 241 |TCATATTTTGGAAACAGACAGATCACCGGTATCAACATAGCTGGATGGAACAGATGGC 300  
Db 365 |TCATATTTTGGAAACAGACAGATCACCGGTATCAACATAGCTGGATGGAACAGATGGC 424  
Qy 301 |AAATGGCAGAGACATGTGCAATGTGTGCTCAGCTGGAAGATATAAATTTGCAATTAAT 360  
Db 425 |AAATGGCAGAGACATGTGCAATGTGTGCTCAGCTGGAAGATATAAATTTGCAATTAAT 484  
Qy 361 |TTAGATTGGATCAACGGTTACCGGAATACATCATCTGTGTAACAATCTCCGTGTGTT 420  
Db 485 |TTAGATTGGATCAACGGTTACCGGAATACATCATCTGTGTAACAATCTCCGTGTGTT 544

421 GGTGCTGAGGTGGCTTATTTATTTGATGTTCTCATGAAAAATTTGAATATTCCTCTTCT 480  
545 GGTGCTGAGGTGGCTTATTTATTTGATGTTCTCATGAAAAATTTGAATATTCCTCTTCT 604  
481 AAAGTGCACTTGATTTGGCCACAGCTTGGGRCACACCTGCTGGGAGAGCTGGTCAAGG 540  
605 AAAGTGCACTTGATTTGGCCACAGCTTGGGAGCACACCTGCTGGGAGAGCTGGTCAAGG 664  
541 ATACCAGGCTTTGGAAGAAATAACTGGGTGGACCCAGCTGGGCCATTTTCCACAACACT 600  
665 ATACCAGGCTTTGGAAGAAATAACTGGGTGGACCCAGCTGGGCCATTTTCCACAACACT 724  
601 CCAAGGAGCTCAGCTAGACCCCTCGATGCCAACTTTGTTGACGTTATTCATACAAAAT 660  
725 CCAAGGAGCTCAGCTAGACCCCTCGATGCCAACTTTGTTGACGTTATTCATACAAAAT 784  
661 GCAGCTCGATCCTCTTTGAGCTTGGTGTGGAAACCAATGATGCTTGCTGTCATCTTGAC 720  
785 GCAGCTCGATCCTCTTTGAGCTTGGTGTGGAAACCAATGATGCTTGCTGTCATCTTGAC 844  
721 TTTTACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
845 TTTTACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 904  
781 AAATTTAACTTCAATGCTTACAAAAGAAAGAAATGCTTCTTTGACGTTAATTAACCTTAC 840  
905 AAATTTAACTTCAATGCTTACAAAAGAAAGAAATGCTTCTTTGACGTTAATTAACCTTAC 964  
841 CGAAGTTATCAATTTATGCTGGAAGCACTTCTTAATCCTGATGCAATTTATGCTTATCCT 900  
965 CGAAGTTATCAATTTATGCTGGAAGCACTTCTTAATCCTGATGCAATTTATGCTTATCCT 1024  
901 TGTAGATCCTACATCTTTTAAAGCAGGAAATGCTTCTTTGCTGAGCAATTTATGCTTATCCT 960  
1025 TGTAGATCCTACATCTTTTAAAGCAGGAAATGCTTCTTTGCTGAGCAATTTATGCTTATCCT 1084  
961 CCAACAAATGGCTGCTATTTGCTGATAGATTTTCACTTCAAAAATATGAAGCTAATGATCA 1020  
1085 CCAACAAATGGCTGCTATTTGCTGATAGATTTTCACTTCAAAAATATGAAGCTAATGATCA 1144  
1021 CATTAATTTTAAACACAGGCTCCTTTCCCAATTTGCCGTTGGAGGACAAATTTGCT 1080  
1145 CATTAATTTTAAACACAGGCTCCTTTCCCAATTTGCCGTTGGAGGACAAATTTGCT 1204  
1081 GTTAACTCAGTGAAGGAGTCACTCAAGGAGCTGCTTCTTCTGCTGAGCGGGCA 1140  
1205 GTTAACTCAGTGAAGGAGTCACTCAAGGAGCTGCTTCTTCTGCTGAGCGGGCA 1264  
1141 ATGGGAAACTGGGGAGTTTGCATTTGTCAGTGGAAACTTGGAGCCAGGCACTTAC 1200  
1265 ATGGGAAACTGGGGAGTTTGCATTTGTCAGTGGAAACTTGGAGCCAGGCACTTAC 1324  
1201 ACAGAAATTAATCGATGAGATGTTAAAGTTGGAAACATTAACAGTGTTCAGTTCTGG 1260  
1325 ACAGAAATTAATCGATGAGATGTTAAAGTTGGAAACATTAACAGTGTTCAGTTCTGG 1384  
1261 AAAAAACATTTGTTTGAAGATTCTCAGAAATAGTTGGAGCAGAAATGTTGATAATACA 1320  
1385 AAAAAACATTTGTTTGAAGATTCTCAGAAATAGTTGGAGCAGAAATGTTGATAATACA 1444  
1321 TCTGGAAATATGATATAAATCTAATCTTCTGAGGCAAGACATTAATGGAACCTAATAT 1380  
1445 TCTGGAAATATGATATAAATCTAATCTTCTGAGGCAAGACATTAATGGAACCTAATAT 1504  
1381 CTCAGAACCTGAAACCATGCTAA 1404  
1505 CTCAGAACCTGAAACCATGCTAA 1528

Publication No. US20030165975A1  
GENERAL INFORMATION:  
APPLICANT: Khodadoust, Mehran  
APPLICANT: Kapeller-Libermann, Rosana  
TITLE OF INVENTION: No. US20030165975A1el Human Lipase Proteins, Nucleic Acids Encodi  
FILE OF INVENTION: Then, and Uses of Both of These  
FILE REFERENCE: 10147-14  
CURRENT APPLICATION NUMBER: US/10/403,745  
CURRENT FILING DATE: 2003-03-31  
PRIOR APPLICATION NUMBER: US/09/411,132A  
PRIOR FILING DATE: 2000-09-12  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 2  
LENGTH: 1401  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-403-745-2

Query Match 99.7%; Score 1399.4; DB 13; Length 1401;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGCTTGAATTTGGATTGTTGTCATCTTCTTTGGCACATCAAGAGGAAAAAGAGTT 60  
Db 1 ATGCTTGAATTTGGATTGTTGTCATCTTCTTTGGCACATCAAGAGGAAAAAGAGTT 60  
QY 61 TGCTATGAAAGGTTAGGCTGTTTCAAAGATGTTTACCAGGACCAAGGACTTTCTCAACA 120  
Db 61 TGCTATGAAAGGTTAGGCTGTTTCAAAGATGTTTACCAGGACCAAGGACTTTCTCAACA 120  
QY 121 GAGTGTGATGTTTACCTGGTCTCCAGAGAGATAAACACTGCTGTTTCTGCTTACACT 180  
Db 121 GAGTGTGATGTTTACCTGGTCTCCAGAGAGATAAACACTGCTGTTTCTGCTTACACT 180  
QY 181 ATACACAATCCCAATGCTTATCAGGAGATCAGTGGGTTAAATTTTCAACTATCCAAGCC 240  
Db 181 ATACACAATCCCAATGCTTATCAGGAGATCAGTGGGTTAAATTTTCAACTATCCAAGCC 240  
QY 241 TCATATTTTGGAAACAGACAAGATCAACCGTATCAACATAGCTGGATGGAAAAACAGATGGC 300  
Db 241 TCATATTTTGGAAACAGACAAGATCAACCGTATCAACATAGCTGGATGGAAAAACAGATGGC 300  
QY 301 AAATGGCAGAGACATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 360  
Db 301 AAATGGCAGAGACATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 360  
QY 361 TTAGATTGGATCAACCGTTTACGGGAAATATCCATGCTGTAAACAATCTCCGCTGTTGT 420  
Db 361 TTAGATTGGATCAACCGTTTACGGGAAATATCCATGCTGTAAACAATCTCCGCTGTTGT 420  
QY 421 GGTGCTGAGTGGCTTATTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTT 480  
Db 421 GGTGCTGAGTGGCTTATTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTT 480  
QY 481 AAAGTGCACTTGATTTGGCCACAGCTTGGGAGCACACCTGCTGGGAGAGCTGGGTCAAGG 540  
Db 481 AAAGTGCACTTGATTTGGCCACAGCTTGGGAGCACACCTGCTGGGAGAGCTGGGTCAAGG 540  
QY 541 ATACCAGGCTTTGGAAGAAATAACTGGGTGGACCCAGCTGGGCCATTTTCCACAACACT 600  
Db 541 ATACCAGGCTTTGGAAGAAATAACTGGGTGGACCCAGCTGGGCCATTTTCCACAACACT 600  
QY 601 CCAAGGAGCTCAGGCTAGACCCCTCGATGCCAACTTTGTTGACGTTATTCATACAAAAT 660  
Db 601 CCAAGGAGCTCAGGCTAGACCCCTCGATGCCAACTTTGTTGACGTTATTCATACAAAAT 660  
QY 661 GCAGCTCGATCCTCTTTGAGCTTGGTGTGGAAACCAATGATGCTTGCTGTCATCTTGAC 720  
Db 661 GCAGCTCGATCCTCTTTGAGCTTGGTGTGGAAACCAATGATGCTTGCTGTCATCTTGAC 720  
QY 721 TTTTACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780

721 TTTTACCAATGAGGAGGACATGCCAGGATGTGAAGACTTAATTACACCTTTACTG 780  
 781 AAATTTAACTTCAATGCTTCAAAAAGAAATGGCTTCTTTTACACTGAACCATGCC 840  
 781 AAATTTAACTTCAATGCTTCAAAAAGAAATGGCTTCTTTTACACTGAACCATGCC 840  
 841 CGAAGTTATCAATTTTATGCTGAAGCAATTTCTTAATCCTGATGCAATTTATGCTTATCCT 900  
 841 CGAAGTTATCAATTTTATGCTGAAGCAATTTCTTAATCCTGATGCAATTTATGCTTATCCT 900  
 901 TGTAGATCCTACATCTTTTAAAGCAGGAAATGCTTTTGTTCCTCAAGAGAGGTTGC 960  
 901 TGTAGATCCTACATCTTTTAAAGCAGGAAATGCTTTTGTTCCTCAAGAGAGGTTGC 960  
 961 CCAACAATGGGTCAATTTTGTCTGATAGATTTCACTTCAAAAATATGAAGACTTAATGATCA 1020  
 961 CCAACAATGGGTCAATTTTGTCTGATAGATTTCACTTCAAAAATATGAAGACTTAATGATCA 1020  
 1021 CATTATTTTAAACACAGGGTCCCTTTCCCAATTTTCCCGTTTGGAGGCACAAATTTGTCT 1080  
 1021 CATTATTTTAAACACAGGGTCCCTTTCCCAATTTTCCCGTTTGGAGGCACAAATTTGTCT 1080  
 1081 GTTAAATCTCAGTGAAGGAGTCACTCAAGGAACTGTCTTTCTTCTGTAGCGGGGCA 1140  
 1081 GTTAAATCTCAGTGAAGGAGTCACTCAAGGAACTGTCTTTCTTCTGTAGCGGGGCA 1140  
 1141 ATTGGGAAACTGGGGAGTTGGCAATTTGTCTGATGAGGAACTTGAAGGCAATGATTTAC 1200  
 1141 ATTGGGAAACTGGGGAGTTGGCAATTTGTCTGATGAGGAACTTGAAGGCAATGATTTAC 1200  
 1201 ACAAATTTAATCGATGAGATGTTAAGTTTGGAAACATTACAGGTTTCAGTTCTCTCG 1260  
 1201 ACAAATTTAATCGATGAGATGTTAAGTTTGGAAACATTACAGGTTTCAGTTCTCTCG 1260  
 1261 AAAAAACATTTGTTGAAGATTCTCAGAATAAGTTGGGAGCAAAATGGTGATAATACA 1320  
 1261 AAAAAACATTTGTTGAAGATTCTCAGAATAAGTTGGGAGCAAAATGGTGATAATACA 1320  
 1321 TCTGGGAAATATGATATAATCTACCTTCTGTAGGCAAGCAATTTAGGACCTTAATTT 1380  
 1321 TCTGGGAAATATGATATAATCTACCTTCTGTAGGCAAGCAATTTAGGACCTTAATTT 1380  
 1381 CTCCAGAACCTGAAACCATGC 1401  
 1381 CTCCAGAACCTGAAACCATGC 1401

US-10-038-1  
 Application No. US20030219862A1  
 PUBLICATION INFORMATION:  
 PUBLICANT: Agarwal, Pankaj  
 PUBLICANT: Cogswell, John P.  
 PUBLICANT: Kabnic, Karen S.  
 PUBLICANT: Lai, Ying-Ta  
 PUBLICANT: Martensen, Shelby A.  
 PUBLICANT: Murdock, Paul R.  
 PUBLICANT: Smith, Randall F.  
 PUBLICANT: Strum, Jay C.  
 PUBLICANT: Xiang, Zhaoying  
 PUBLICANT: Xie, Qing  
 PUBLICANT: Rizni, Safia K.  
 TITLE OF INVENTION: NOVEL COMPOUNDS  
 (LE REFERENCE: GP50029)  
 CURRENT APPLICATION NUMBER: US/10/312,088  
 CURRENT FILING DATE: 2002-12-20  
 PRIOR APPLICATION NUMBER: PCT/US01/19929  
 PRIOR FILING DATE: 2001-06-22  
 PRIOR APPLICATION NUMBER: 60/213,161  
 PRIOR FILING DATE: 2000-06-22  
 PRIOR APPLICATION NUMBER: 60/213,156  
 PRIOR FILING DATE: 2000-06-22

; NUMBER OF SEQ ID NOS: 44  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 1383  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-312-088-1

Query Match 87.9%; Score 1234.6; DB 13; Length 1383;  
 Best Local Similarity 93.8%; Pred. No. 0;  
 Matches 1314; Conservative 0; Mismatches 69; Indels 18; Gaps 2;  
 Qy 1 ATGCTTGGAAATTTGGATTTGCTTCTTTTGGCAGATCAAGAGGAAAGAGTT 60  
 Db 1 ATGCTTGGAAATTTGGATTTGCTTCTTTTGGCAGATCAAGAGGAAAGAGTT 60  
 Qy 61 TGTATGAAAGTTAGGTTTCAAGATGTTTACCATGAGCAGGACTTTCACACA 120  
 Db 61 TGTATGAAAGTTAGGTTTCAAGATGTTTACCATGAGCAGGACTTTCACACA 120  
 Qy 121 GAGTTGGTAGGTTTACCTTGGTCTCCAGAGAGATAAACAACCTCGTTTCTGCTTACACT 180  
 Db 121 GAGTTGGTAGGTTTACCTTGGTCTCCAGAGAGATAAACAACCTCGTTTCTGCTTACACT 180  
 Qy 181 ATACACAATCCCAATGCCCTATCAGGAGATCAGTGGGTTAATTTCAACTATCAAGCC 240  
 Db 181 ATACACAATCCCAATGCCCTATCAGGAGATCAGTGGGTTAATTTCAACTATCAAGCC 240  
 Qy 241 TCATATTTTGGAAACAGAGATCAACCGTATCAACATAGCTGGATGGAACAGATGGC 300  
 Db 241 TCATATTTTGGAAACAGAGATCAACCGTATCAACATAGCTGGATGGAACAGATGGC 300  
 Qy 301 AAATGCGAGAGACATGTCAAATGTTTGTCTACAGCTGGAAGATATAAATTTGCATTAA 360  
 Db 301 AAATGCGAGAGACATGTCAAATGTTTGTCTACAGCTGGAAGATATAAATTTGCATTAA 360  
 Qy 361 TTAGTTGATCAAGCGTTTCAAGGATACATCCATGCTGTAACCAATCTCCGTTGTT 420  
 Db 361 TTAGTTGATCAAGCGTTTCAAGGATACATCCATGCTGTAACCAATCTCCGTTGTT 420  
 Qy 421 GGTGCTGAGTGGCTTATTTTATGATGTTTCTCATGAAAAAATTTGAATATTTCCCTTCT 480  
 Db 421 GGTGCTGAGTGGCTTATTTTATGATGTTTCTCATGAAAAAATTTGAATATTTCCCTTCT 480  
 Qy 481 AAAGTGCACCTTGAATTTGGCCACAGCTTGGAGAGACACCTTGGCTGGGAGCTGGGTCAGG 540  
 Db 481 AAAGTGCACCTTGAATTTGGCCACAGCTTGGAGAGACACCTTGGCTGGGAGCTGGGTCAGG 540  
 Qy 541 ATACAGAGGCTTTGGAAGAAATAACTGGGTTGGACCCAGCTGGGCCATTTTCCACAACACT 600  
 Db 541 ATACAGAGGCTTTGGAAGAAATAACTGGGTTGGACCCAGCTGGGCCATTTTCCACAACACT 600  
 Qy 601 CCAAGGAAAGTCAAGGCTAGACCCCTCGGATGCGCAATTTTGTGAGCTGATTTATCATACAA 660  
 Db 601 CCAAGGAAAGTCAAGGCTAGACCCCTCGGATGCGCAATTTTGTGAGCTGATTTATCATACAA 660  
 Qy 661 GCAGCTCGCATCTCTTTGAGCTTGGTGTGGACCAATTTGATGCTTGGTGGTCACTTGCAC 720  
 Db 661 GCAGCTCGCATCTCTTTGAGCTTGGTGTGGACCAATTTGATGCTTGGTGGTCACTTGCAC 720  
 Qy 721 TTTTACCCAAATGGAGGAAAGCACATGCCAGGATGTGAAGACTTAATTTACACCTTTTACTG 780  
 Db 721 TTTTACCCAAATGGAGGAAAGCACATGCCAGGATGTGAAGACTTAATTTACACCTTTTACTG 780  
 Qy 781 AAATTTAACTTCAATGCTTCAAAAAGAAATGGCTTCTTTTGTAGCTGAACCATGCC 840  
 Db 781 AAATTTAACTTCAATGCTTCAAAAAGAAATGGCTTCTTTTGTAGCTGAACCATGCC 840  
 Qy 841 CGAAGTTATCAATTTTATGCTGAAAGCAATTTCTTAATCCTGATGCAATTTATTTGCTTATCCT 900  
 Db 841 CGAAGTTATCAATTTTATGCTGAAAGCAATTTCTTAATCCTGATGCAATTTATTTGCTTATCCT 900  
 Qy 901 TGTAGATCCTACATCTTTTAAAGCAGGAAATTTGCTTCTTTTGTTCACAAAGAGGTTGC 960

901 |||||TGATATCCATACATCTTTTAAAGCAGTACATGTGTAGGATGTGC-----AGATTT 953  
961 CCAACAATGGGTCAATTTCTGATAGATTTCACTTCAAAATATGAAGACTAATGGATCA 1020  
954 GTTACATAGATAGATAGATAGATAGGAGTACATCTCCATGTGTTTTAACCCCTTCTCT 1013  
1021 CATTAATTTTAAACACAGGGTCCCTTTCCCATTTGGCCGTTGGAGGCACAAATTTGCT 1080  
1014 CCGTTTCCCTT-----CTTGTTCCTTATATCTAGGTGTGGAGGCACAAATTTGCT 1062  
1081 GTTAAACTCASTGGGAAGCAAGTCACTCAAGGAAGTCTCTTTCTGTTAGGCGGGCA 1140  
1063 GTTAAACTCAGTGGGAAGCAAGTCACTCAGGAAGTCTCTTTCTGTTAGGCGGGCA 1122  
1141 ATTGGAAAACCTGGGAGTTTCCATTTGCTAGTGGAAAACCTTGAGCCAGGATGACTTAC 1200  
1123 GTTGGAAAACCTGGGAGTTTCCATTTGCTAGTGGAAAACCTTGAGCCAGGATGACTTAC 1182  
1201 ACAAATTAATCGATGCAGATGTTACCTTGGAAACATTAAGTGTTCAGTTCACTCTGG 1260  
1183 ACAAATTAATCGATGCAGATGTTACCTTGGAAACATTAAGTGTTCAGTTCACTCTGG 1242  
1261 AAAAAACATTTTGAAGATTTCTCAGATTAAGTTGGGAGCAGAAATGGTGATAATACA 1320  
1243 AAAAAACATTTTGAAGATTTCTCAGATTAAGTTGGGAGCAGAAATGGTGATAATACA 1302  
1321 TCTGGGAATATGGATATAAATCTACCTTCTGTAGCCAAAGACATATTGGGACCTTAATTT 1380  
1303 TCTGGGAATATGGATATAAATCTACCTTCTGTAGCCAAAGACATATTGGGACCTTAATTT 1362  
1381 CTCAGAACCTGAACCATGC 1401  
1363 CTCAGAACCTGAACCATGC 1383

T 5  
-969-347-193  
ence 191, Application US/09969347  
ent No. US20020115085A1  
PRAL INFORMATION:  
PLICANT: Ebner, Reinhard  
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu  
LE OF INVENTION: Sets  
LE REFERENCE: 689290-69  
RENT APPLICATION NUMBER: US/09/969,347  
RENT FILING DATE: 2001-10-02  
FOR APPLICATION NUMBER: US/60/237,598  
FOR FILING DATE: 2000-10-03  
FOR APPLICATION NUMBER: US/60/237,604  
FOR FILING DATE: 2000-10-03  
MBER OF SEQ ID NOS: 318  
FTWARE: PatentIn version 3.0  
ID NO 193  
ENGTH: 1481  
YPE: DNA  
RGANISM: Homo sapiens  
-969-347-193

ry Match 27.5%; Score 386.2; DB 10; Length 1481;  
t-Local Similarity 57.1%; Pred. No. 3e-101;  
ches 807; Conservative 0; Mismatches 589; Indels 18; Gaps 5;

1 ATGCTTGGAAATTTGGATTTGTCATTTCTTTGTCATCATCAAGAGGAAAGAGATT 60  
20 ATGCTGATCTTCTGGACATATCACAATTTCTCTGCTGGGAGCAGCAAGGAAAGAGATT 79  
61 TCGTATGAAGGTTAGGGTGTTCAGAGATGTTTACATGGACAGGACTTCTCAACA 120  
80 TCGTATGAGGACCTCGGGTGTCTTCTGACATGAGCCCTGGGCGGGACAGCAATCAGG 139  
121 GAGTTGGTAGTTTACCTCGGTCTCCAGAGAAGATTAACATCTCGTTTCTGCTCTACACT 180

140 CCCCTGAAATTTCTCCCTGGAGCCCTGAGAAGATCGGCACCCCGCTTCTGCTGTACACC 199  
181 ATACACAATCCCAATGCCCTATCAGAGATCAGTGGGTAAATTTCTCAATCTATCCAGGCC 240  
200 AATGAAACCCAAACAACTTTCAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 259  
241 TCATATTTTGAACACAGACATCAACCGTATCAACATAGCTGG-----ATGGAACA 294  
260 TCAATTTTCAATGGACAGAAAGACCCGGTTCATCATCCATGGCTTTCATAGACAAAGGA 319  
295 GATGCAATGGCAGAGAGACATGTGCAATGTGTGTCTACAGCTGGAGATATAAATTC 354  
320 GATGAGACTGGGTGACAGACATGTGCAAGAACTGTTTCGAGGTGGAGAGGTGAATGC 379  
355 ATTAATTTAGATTTGGATCAACGGTTTCACGGG---AATACATCCTCATGTGTAAACAATCTC 411  
380 ATCTCGTGGACTGGAAGAGGGCTCCCAAGCCACCTACACACAGGCTGCCAACACAGTG 439  
412 CGTGTGTTGTTGCTGAGGTGCTTATTTATTTGATGTTTCTCATGAAAAAATTTGAATAT 471  
440 CGAGTGGTGGGCGGCCAGGTGGCCAGATGCTGCATCTCTTTGACAGAGTATAGCTAC 499  
472 TCCCTCTTAAAGTGCATTTGATTTGGCCACAGCTTTGGAGCACACCTGGCTGGGGAAGCT 531  
500 CCCCCCTTCCAAAGTTCACTCATTTGGCCACAGCTTGGAGCCACAGTGGCTGGAGAGCA 559  
532 GGGTCAAGGATACCAAGCCCTTGAAGAAATTAATCTGGGTTGGACCCAGCTGGGCCATTTTTC 591  
560 GGAAGCAAGACTCTCAGGCTTGAAGAGGATTTACAGGTTGGATCTCTGTAGAAGCAAGTTTC 619  
592 CACACACTCTCAAAAGGAAGTCAAGCTAGACCTGACCTCGGATGCAACTTTGTTGAGCTTAT 651  
620 GAGAGTACTCTGAAGAGGTGCGACTTGTATCCCTCTGATGTGACTTTTGTGATGTGATT 679  
652 CATACAAATGACAGCTCGCATCTCTTTGAGCTTGGTGTGGAACCAATGATGCTTGTGGT 711  
680 CACAGGATGAGCTCCCTGATCCCATTTCTTGGTTTGGAAACCAACACAGATGGGT 739  
712 CATCTTGACTTTTACCCAAATGAGGAGGAGACACATGCCAGGATGTGAGACTTAATTAACA 771  
740 CATCTTGACTTTTCCCAATGAGGAGGAGAGCATGCCGGATGCAAGAA---GAATGCC 796  
772 CTTTACTGAAATTTAACTTCAATGCTTTACAAAAGAAATGGCTTCTTCTTTTGAAGT 831  
797 CTGCTCAGATCGTGGATCTAGATGGCATCTGGSCGGAAACCCGGGACTTTTGGCTTGC 856  
832 AACATGCCCGAAGTTATCAATTTTATGCTGAAAGCAATCTTAATCTGATGCAATTTAT 891  
857 AATCACCTTAAGAAGCTACAAGTATTAATTTGGAAGAGCATCTCAATCCGATGGGTTGCT 916  
892 GCTTATCTTGTAGATCTACACATCTTTTAAAGCAGGAAATGCTTCTTTTGTCCAAA 951  
917 GCATATCCCTGCACTTCTTACAGTCTTTTGAAGTCTGACAGTGTCTCCCGTGTCCAGAT 976  
952 GAAGTTGCCCAACAATGGGTCAATTTTGTGATAGATTTCACTTTCAAAATATGAAGACT 1011  
977 CAAGGATGCCCAAGATGGGTCACTATGCTGATAAATTTGCTGCGCAGGACAAGTGA---A 1033  
1012 AATGATCACAATTTTAAACACAGGCTCCCTTTCCCAATTTGCCGTTGGAGGCAC 1071  
1034 GAGCAGCAAAATTTCTTCTTGAACACAGGAGAGCTAGCAATTTGCTCTGCTGGAGATAT 1093  
1072 AAATTTGCTGTAAACTCAGTGGAGGCAAGTCACTCAAGGAACTGTCTTTTCTTCTGTTA 1131  
1094 GGGGTTTCCATCACAATCTGTGGAAGAAACAGCCACT---GGTCAGATCAAAAGTTGCTTTG 1150  
1132 GCGGGGCAATTTGGAAAACCTGGGAGTTTGCATTTGCTAGTGGAAAACCTTGAGCCAGGC 1191  
1151 TTTGGAATAAGGAAACACTCACAGTACAGCATCTTCAGGGGATTTCTCAAAACCCAGGC 1210  
1192 ATGACTTACAAAAATTAATGATGAGATGTTTAACTGTTGAAAACATTAAGTGTTCAG 1251  
1211 TCACCCATCTATGATGTTGATGCAAGCTGGATGTTGGAACATTTGGAAGTCAAG 1270

1252 TTCTCTGGAAGAAACATTTGTTGAAGATTCTCAGAAATAGTTGGAGCAGAAATGGTG 1311  
|||||  
1271 TTCTCTGGAATACAAATGTGATAAATCCAACTCTCCCAAGTGGTGCCCAAGATC 1330  
|||||  
1312 ATAATAATCTGGGAATATGGATATAAATCTACCTTCTGTAGCCAGACATATGGGA 1371  
|||||  
1331 ACTGTGCAAAAGGAGAGAGAGACAGTGTACAACTTCTGTAGCGAAGACACAGTGGG 1390  
|||||  
1372 COTATATCTCCAGAACTGAAACCATGCTAA 1404  
|||||  
1391 GAAGACAGCTGCTCACCTCACCTCACCTGCTAA 1423  
|||||

## LT 6

9-923-779-145  
quence 145, Application US/09923779  
tent No. US20020076721A1  
NERAL INFORMATION:  
PLICANT: Pyle, Ruth A.  
PLICANT: Xu, Jiangchun  
PLICANT: Kalos, Michael D.  
ILE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
ILE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER  
ILE REFERENCE: 210121.553  
URRENT APPLICATION NUMBER: US/09/923,779  
URRENT FILING DATE: 2001-08-06  
UMBER OF SEQ ID NOS: 155  
FTWARE: FastSeq for Windows Version 4.0  
Q ID NO 145  
LENGTH: 1471  
TYPE: DNA  
ORGANISM: Homo sapiens  
9-923-779-145

ery Match 24.4%; Score 342.6; DB 9; Length 1471;  
st Local Similarity 57.0%; Pred. No. 1.5e-88;  
tches 791; Conservative 0; Mismatches 569; Indels 27; Gaps 8;  
29 TGTCTTTGGCAGATCAGAGGAAAGATTTGCTATGAAAGTTAGGGTGTTCAGAG 88  
38 TGCTGTGGGAGCAGTAGCAGGAAAGATTTGCTACGAAAGACTCGGCTGCTCAGTG 97  
89 ATGGTTTACCATGACCAGGACTTTCTCAACAGAGTTGGTAGGTTTACCCTGGTCTCCAG 148  
98 ATGACTCCCATGTCAGGAATACGGAAGAGCCCTCCATATATTGCTTGGTCTCCGA 157  
149 AGAAGATAAAGACTCGTTTCTGCTCTACACTATACAAATCCCAATGCCATATCAGGAGA 208  
158 AAGATGTCAACACCCGCTTCTCTATATATCTAATGAGAACCCAAACAACTTTCAAGAAG 217  
209 TCAGTGGGGTTAATCTTTCAACTATCCAGCCTCATATTTTGGAAACAGACAGATCACCC 268  
218 T---TGCGCAGATTCTAAGCATCAGTGGCTCCAAATTTCAAAACAAATAGAAAATC 274  
269 GTATCAACATAGCTGGAT-----GGAAAAAGATGGCAATGGCAGAGACATGTGCA 322  
275 GCTTTATTATTATGATGATTCATAGACAAAGGAGAGAAAGAACTGGCTGGCAATGTGCA 334  
323 ATGTGTTGTCTACAGCTGGAAGATATAAATTTGCAATTAATTAGATTGGATCAAGCTTAC 382  
335 AGAATCTGTTCAAGTGGGAAGTGTGAACTGTATCTGTGTGGACTGGAAAGTGGCTCCC 394  
383 GG---GAATACATCCATGCTGTATAACAAATCCCGTGTGTTGTTGCTGAGGTGGCTTATT 439  
395 GAATCGTATACACAGAGCTCGCAGAACATCAGGATCGTGGGAGCAGAGTGGCATATT 454  
440 TTATTGATGTTCTCATGAAATAATTTGAATATTTCCCTTCTAAAGTGCACTGATGGCC 499  
455 TTGTGTAATTCTTTCAGTCCGCGTTCGGTTACTACCTTCCAAAGTGTGATGTCATTGCC 514  
500 ACAGCTTGGGAGCACACTCGCTGGGAGAGCTGGGTCAAGGATACAGG---CCTTGGAA 556  
|||||

Db 515 ACACCTGGGTGCCACGCTGCTGGGAGGCTGGAAGGAGAACCAATGGGACCATTTGGAC 574  
Qy 557 GAATACTGGTTGGACCCAGCTGGGCCATTTTCCACAACTCCAAAGGAAGTCAGGC 616  
Db 575 GCATCAGAGGTTGGACCCAGCAGAACCTTGTCTTCAGGGCAGACCTGAATTAGTCCGAT 634  
Qy 617 TAGACCCCTCGGATGCAAACTTTGTTGAGCTTATTATCATACAAATGAGCTCGCATCCTCT 676  
Db 635 TGGACCCACCGATGCCAAATTTGTGATGTAATTCACACGGATGTTGCCCCCATAGTCC 694  
Qy 677 TTGAGCTTGTGTTGGAAACCAATTTGATGCTGTGTGTCATCTTGACTTTTACCCAAATGGAG 736  
Db 695 CCAATTTGGGTTTGAATGAGCCAAAGTCGTGGGCCACCTAGATTCTTTCCAAATGGAG 754  
Qy 737 GGAAGCACATGCCAGGATGTGAAGACTTAATACACCTTTACTGAAATTTAACTTCAATG 796  
Db 755 GAGTGGAAATGCTGGATGTAAAGAACATT---CTCTCTCAGATTGTGGACATAGAG 811  
Qy 797 CTTACAAAAGAAATGGCTTCTCTTTGACTGTAAACATGCCCGGAGTTATCAATTTT 856  
Db 812 GAATCTGGGAAGGAGACTCGAGACTTTTGGCGCCTGTAACTTAAGAAAGCTACAAATAT 871  
Qy 857 ATGCTGAAGCACTTTAATCTCTGATGCAATTTATTGCTTATCTTGTAGATCCTACACAT 916  
Db 872 ACATGATGATCGTCAACCCCTGATGGCTTTGCTGGATTCCCTGTGCTCTTACAAAG 931  
Qy 917 CTTTAAAGCAGGAAATGCTTCTTTTGTTCAAAGAAAGTTTGCACAAACAAATGGTCAAT 976  
Db 932 TCTTCACTGCAAAAGATGTTTCCCTTGTCCAAGTGGAGCTGCCACAGATGGGTCACT 991  
Qy 977 TTGCTGATGATTTCACTTCAAAATATGAAGACTAAATGGATCACTATTATTTTAAACA 1036  
Db 992 ATGCTGATGATATCTCTGGGAAA---ACAAATGATGTGGGCCAGAAATTTTATCTAGACA 1048  
Qy 1037 CAGGGTCCCTTTCCCAATTTGCCCGTTGGAGGCACAAATTTGCTTTAAACTCAGTGGAA 1096  
Db 1049 CTGGTATGTCAGTAATTTTTCAGGTTGGAGGTATAAGGTATCTGTCACTGTCTGGAA 1108  
Qy 1097 GCAGATCACTCAGGAACTGTCTTCTCGTGTAGGGGGCAATTCGGHAAACTGGGG 1156  
Db 1109 AAAAGGTAC---AGGACACACTACTAGTCTTTCTTTCGGAAATTAAGGAAACTCTAAGC 1165  
Qy 1157 AGTTTCCCATTTGTCAGTGGAAACTTTAGCCAGGAGCATCTATACAAAAATTAATCGATG 1216  
Db 1166 AGTATGAATTTTCAAGGCACTCTCAACACAGATAGTACTCATTTCCATGAAATTTGACT 1225  
Qy 1217 CAGATGTTAAGTGGAAACATTAACAGTGTTCAGTTCATCTCGGAAAAAACAATTTGTTG 1276  
Db 1226 CAGATGTGATGTTGGGACTTGCAGATGGTTAAATTTTATTTGGTATTAACAATGTGATCA 1285  
Qy 1277 AAGATTCTCAGAAATAGTTGGGAGCAGAAATGTTGATAAATACATCTCGGAAATATGGAT 1336  
Db 1286 ACCCACTTTTCTAGTGGGAGCATCCAGATTATAGTGAGAC---AAATGTTGAA 1342  
Qy 1337 ATAAATCTACCTTCTGTAGCAGACATTAATGGGACCTPAATATTCTCCAGAACCTGAAAC 1396  
Db 1343 AACAGTTCAACTTCTGTAGTCCAGAAACCGTCAAGGAGGAAAGTTCTGTCTCACCCTCACAC 1402  
Qy 1397 CATGCTA 1403  
Db 1403 CGTGTTA 1409

## RESULT 7

US-09-969-347-176  
; Sequence 176, Application US/09969347  
; Patent No. US2002011508A1  
; GENERAL INFORMATION:  
; APPLICANT: Ebner, Reinhard  
; TITLE OF INVENTION: Cancer  
; TITLE OF INVENTION: Sets  
; FILE REFERENCE: 689290-69  
; CURRENT APPLICATION NUMBER: US/09/969,347

RENT FILING DATE: 2001-10-02  
 IOR APPLICATION NUMBER: US/60/237,598  
 IOR FILING DATE: 2000-10-03  
 IOR APPLICATION NUMBER: US/60/237,604  
 IOR FILING DATE: 2000-10-03  
 MEMBER OF SEQ ID NOS: 318  
 IOWARE: Patent in version 3.0  
 ID NO 176  
 LENGTH: 1471  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 -969-347-176

ry Match 24.4%; Score 342.6; DB 10; Length 1471;  
 t Local Similarity 57.0%; Pred. No. 1.5e-88;  
 ches 791; Conservative 0; Mismatches 569; Indels 27; Gaps 8;

29 TGTCTCTTGGCAGCATCAAGAGGAAAGAGTTGCTATGAAGGTTAGGGTTTCAAG 88  
 38 TGCTGCTGGAGCAGTAGCAGGAAAGAAAGTTTGTACGAAGACTCGGCTGCTTCAGTG 97  
 89 ATGGTTTACCATGGACAGGACTTTCTCAACAGAGTTGGTAGGTTTACCTGGTCTCCAG 148  
 98 ATGACTCCCATGCTGAGGAAATTACGGAAGACCCTCCATATATTGCTTGGTCTCAA 157  
 149 AGAGATAAACAATCTGTTTCTGCTCTACACTATACAAATCCCAATCCCTATCAGGAGA 208  
 158 AAGATGCAACACCCGCTTCTCTCTATATCTAATGAGAACCCCAACAACCTTCAAGAAG 217  
 209 TCAGTGGGTTAATCTTCAATCTCAAGCCTCATATTTTGGAAACAGACAGATCAACC 268  
 218 T----TGCCCGAGNATCATCAAGCATCAGTGGCTCCAAATTTCAAAACAATGAAGAACTC 274  
 269 GTATCAACATAGCTGAT-----GGAACACAGATGGCAATGGCAGAGACATGTGCA 322  
 275 GCTTTATTATTCATGATTCATAGACAGAGGAGAGAAACTGGCTGGCCAAATGTGTGCA 334  
 323 ATGTGTTGCTCAGCTGGAGATATATAATTTGATTTAGTTGGATCAACGGTTTCAAC 382  
 335 AGAATCTGTTCAAGGTGGAAAGTGTGAATCTGTCTGTGGACTGGAAAGGTGGCTCCC 394  
 383 GG---GAATACATCCATGCTGTAAACAATCTCCGTGTGTTGGTGTCTGAGGTGGCTTATT 439  
 395 GAACCTGGATACACACAAGCCTCGCAGACATCAGGATCGTGGAGCAGAAAGTGGCATATT 454  
 440 TTATTGATGTTCTCATGAAATTTTGAATATTCCTCTTCAAGTGCACCTTGAATGGCC 499  
 455 TTGTTGAATTTCTTCAAGTCCGCTTCCGTTACTCACTTCCAAACGTGATGCTCATGGCC 514  
 500 ACAGCTTGGGAGCACACCTGGCTGGGGAAGCTGGGTCAAGGATACAGG---CCTTGGAA 556  
 515 ACAGCTTGGTGGCCACGCTGTGGGAGGCTGGAGGAGAACCAATGGGACCAATGGAC 574  
 557 GAATACTGGTGGACCCAGCTGGGCAATTTTCCAAACAATCCCAAGGAAGTCAAGC 616  
 575 GCATCACAGGTTGGACCCAGCAGAACCTTGGTTTCAAGGACACCTGATTAATAGTCCGAT 634  
 617 TAGACCCCTCGATGACCAATTTGTGAGTGTATTCATCAAAATGCAGCTGCATCTCT 676  
 635 TGACCCCGAGGATGCAATTTGTGGATGTAAATTCACACGATGGTCCCCCATAGTCC 694  
 677 TTGAGCTTGGTGTGGAACCAATGTATGCTGTGGTCACTCTGACTTTTACCCCAATGGAG 736  
 695 CCAATTGGGTTTGGAAATGAGCAAGTGTGGGCACTAGATTTCTTTCCAAATGGAG 754  
 737 GGAAGCAGATGCCAGATGAGACTTAATTAACCTTTTACTGAAATTAACCTCAATG 796  
 755 GAGTGGAAATGCTGTGATGTAAGAAACAATTT---CTCTCTCAGATTTGGACATAGAG 811  
 797 CTTTACAAAAGAAATAGGCTTCTCTTTGACTGTAAACCATGCCCCGAAGTTATCAATTTT 856  
 812 GAATCTGGGAGGAGCTCGAGACTTTCGAGCTTTCGAGCTGTAACTTAAGAGACTACAAATTT 871

QY 857 ATGCTGAAAGCAATTCCTTAATCTGATGATCAATTTATTGCTTATCTCTGTAGATCTACACAT 916  
 Db 872 AACTGATAGCATCGTCAACCTGATGCTTGTGGATTCCTCTGCTCTTACACAG 931  
 QY 917 CTTTAAAGCAGGAAATGCTTCTTTTCCAAAGAGGTTGCCCAACAATGGGTCAAT 976  
 Db 932 TCTTCACTGCAAAACAAGTGTTCCTCTTGTCCAAAGTGGAGGCTGCCCAAGATGGGTCACT 991  
 QY 977 TTGCTGATAGATTTCACTTTCAAAAATATGAAGACTAATGGATCACAATTTTAAACA 1036  
 Db 992 ATGCTGATAGATATCTCGGAAA--ACAAATGATGGGCCAGAAAATTTTATCTAGACA 1048  
 QY 1037 CAGGTGCTTCCCATTTGCCCTTTGGAGGCAAAATTTGTTTAAACTCAGTGGAA 1096  
 Db 1049 CTGGTATGATCAATTTTCAAGGCTTGAAGGATATAAGGTATCTGTCACTGTCTGGAA 1108  
 QY 1097 GCGAAGTCACTCAAGGACTGCTTCTTCTGCTAGCGGGCAATTTGGGAAACTCGGG 1156  
 Db 1109 AAAGGTTAC--AGGACACATCTAGTCTTTTCTTGGTGGAAATAAAGGAAACTCTAAGC 1165  
 QY 1157 AGTTTGCATTTGTCAGTGGAAACTTTGAGCCAGGATGACTTACACAAAATTTAATCGATG 1216  
 Db 1166 AGTATGAAATTTTCAAGGCACTCTCAACACAGATAGTACTCAATTCATGAAATTTGACT 1225  
 QY 1217 CAGATGTTAAAGTTGGAAACATTAAGAGTGTTCAGTTCATCTGGAAAAACAATTTGTTG 1276  
 Db 1226 CAGATGTTGATGTTGGGACTTGCAGATGGTTAAATTTATTTGGTATTAACAATGTGATCA 1285  
 QY 1277 AAGATTTCTCAGAAATGAAGTTGGAGCAGAAATGGTGAATAATACATCTGGAAATATGAT 1336  
 Db 1286 ACCCAACTTTACCTAGCTGGAGCATCCAGATTTAGTGGAGAC---AAATGTTGAA 1342  
 QY 1337 ATAAATCTACCTTCTGTAGCCAGACATTAATGGACCTAATATTCTCCAGAACCTGAAC 1396  
 Db 1343 AACAGTTCAACTTCTGTAGTCCAGAAACCGTCAGGAGGAAGTTCTGCTCACCCTCACAC 1402  
 QY 1397 CATGCTA 1403  
 Db 1403 CGTGA 1409

RESULT 8  
 US-10-191-997-117  
 ; Sequence 117, Application US/10191997  
 ; Publication No. US20030207834A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Oligos Etc., Inc.  
 ; APPLICANT: DALE, Roderic M. K.  
 ; APPLICANT: ARROW, Amy  
 ; APPLICANT: THOMPSON, Terry  
 ; TITLE OF INVENTION: Oligonucleotide-Containing Pharmacological Compositions And Their  
 ; FILE REFERENCE: 54800-5019  
 ; CURRENT APPLICATION NUMBER: US/10/191,997  
 ; PRIOR FILING DATE: 2002-07-10  
 ; PRIOR APPLICATION NUMBER: US 60/303,820  
 ; PRIOR FILING DATE: 2001-07-10  
 ; NUMBER OF SEQ ID NOS: 132  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 117  
 ; LENGTH: 1471  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: NM\_000936  
 US-10-191-997-117

Query Match 24.4%; Score 342.6; DB 13; Length 1471;  
 Best Local Similarity 57.0%; Pred. No. 1.5e-88;  
 Matches 791; Conservative 0; Mismatches 569; Indels 27; Gaps 8;  
 QY 29 TGTCTTGGCAGCATCAAGAGGAAAGAGTTTCTATGAAGTTAGGTGTTTCAAG 88



38 TGCTGCTGGGACGATAGCAGGAAAGAGTTTCTACGAAAGACTCGGTGCTTCAGTG 97  
89 ATGGTTTACATGGACAGGACTTTCTCAACAGAGTTGGTAGGTTTACCTGCTCTCCAG 148  
98 ATGACTCCCATGGTCAGGAATTACGGAAGACCCCTCCATATATTGCCCTGGTCTCCAA 157  
149 AGAAGATAAACACTCGTTTCCTGCTCTACACTATACAAATCCCAATGCCCTATCAGGAGA 208  
158 AAGATGTCAACACCCGCTTCTCTCTATATATACTAATGAGAACCACAACTTTCAGAGAG 217  
209 TCAGTGGCGTTAATTTCTCACTATCCAGGCTCATATTTTGGNACAGACAGATCACCC 268  
218 T---TGCCGACAGATTCATCAAGCATCATGCTGCTCCAAATTTCAAAACAAATAGAAAACTC 274  
269 GATCAACATAGCTGGAT-----GGAAAAACAGATGGCAAAATGGCAGAGACATGTGCA 322  
275 GCTTTATTATTTCATGGATTATAGACAGGAGAGAAACTGGCTGGCCCAATGTGTGA 334  
323 ATGTGCTGTACAGCTGGAAGATATAAATTCATTAATTTAGATTGGATCAACGGTTTAC 382  
335 AGAATCTGTTCAGGTGGAAGAGTGGAACGTATCTGTGTGGACTGGAAAGGTGGCTCCC 394  
383 GG--GAATACATCATCTCTCTAAACATCTCCGTTGTTGTTGCTGAGGTGGCTTATT 439  
395 GAACCTGGATACACAAAGCTCGCAGAACATCATGATCGTGGGAGCAGAGTGGCAATTT 454  
440 TTAATGATGTTCTCATGAAAAAATTTGAATATCCCTTCTAAAGTGCATCTGATGGCC 499  
455 TTGTTGAATTTCTTCAGTGGCGGTTCGGTTACTCACCTTCCAACTGCTCATTTGGCC 514  
500 ACAGCTTGGGAGCACACCTGGCTGGGAGCTGGTCAAGGATACACAGG---CCTTGGAA 556  
515 ACAGCTGGTGGCCACGCTGCTGGGAGGCTGGAAGAGAACCAATGGACCATTTGGAG 574  
557 GAATAACTGGGTGGACCCAGCTGGGCCATTTTTCACAACTCTCCAAAGGAGTCAAGC 616  
575 GCATCACAGGCTGGACCCAGCAGAACCTTGCTTTCAGGACACACCTGAATTAGTCCGAT 634  
617 TAGACCCCTCGATGGCCCAACTTTCTGAGTTTATCATCAAAATGAGCTCGATCCTT 676  
635 TGGACCCAGGATGCCAAATTTGTGATGTAATTCACGGATGGTGGCCCATAGTCC 694  
677 TTGAGCTTGGTGGACCACTGATGCTTGTGCTCATCTTGACTTTTACCCAAATGGAG 736  
695 CCAATTTGGGTTTGGATGAGCCAGCTGCTGGGCCACCTAGATTTCTTCCAAATGGAG 754  
737 GGAAGCATCCGAGGATGGAAGACTTAATACCTTTACTTGAATTTAACTTCAATG 796  
755 GAGTGGAAATCCCTGGATGTAAAGAAACATTT---CTCTCTCAGATTGGACATAGAG 811  
797 CTTACAAAAAAGAAATGGCTTCTTCTGACTGTAAACCATGCCGAGTTTATCAATTTT 856  
812 GAATCTGGAGGAGACTCGAGACTTTCCGCTGTAATCACTTAAAGACTTACAAATTT 871  
857 ATGCTGAAGCAATCTTAATCTGATGCAATTAATGCTTATCTTGTAGATCCTACACAT 916  
872 ACACGTATGATCGTCAACCCCTGATGGCTTGTGATGCCCTGTGCTCTTACAAG 931  
917 CTTTAAAGCAGGAAATTTGCTTTCTTTTCCAAAGAGGTTGCCCAACAAATGGGTCAT 976  
932 TCTTCACTGCAACAGATGTTTCCCTTGTCCAAGTGGAGGCTGCCACAGATGGGTCACT 991  
977 TTGCTGATPAGATTTCACTTCAAAATATGAAGACTAATGGATCACAATTTTAAACA 1036  
992 ATGCTGATPAGATATCTCTGGGAAA---ACAAATGATGGGCCAGAAATTTTATCTAGACA 1048  
1037 CAGGTCCTTTCCCAATTTCCGCTGGGAGCACAATTTGCTGTAACTCACTGAGTGGAA 1096  
1049 CTGTGATGCGAGTAATTTTGCAGTTGGAGGTATAGGTTATCTGTCACTGTCTGGAA 1108  
1097 GCGAAGTCACATCAAGGAACATGCTTCTTCTGCTGAGCGGGGCAATGGGAAAACCTGGG 1156  
1109 AAAAGGTTAC---AGGACACATACATAGTTCTTCTTGTTCGGAAATTAAGGAACCTTAGC 1165

RESULT 9  
US-09-925-297-108  
; Sequence 108, Application US/09925297  
; Patent No. US20020081659A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA105  
; CURRENT APPLICATION NUMBER: US/09/925,297  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05989  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 928  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 108  
; LENGTH: 1536  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1482)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-297-108

Query Match 24.4%; Score 342.6; DB 9; Length 1536;  
Best Local Similarity 57.0%; Pred. No. 1.5e-88;  
Matches 791; Conservative 0; Mismatch 569; Indels 27; Gaps 8;  
QY 29 TGTTCTTTGGCAGCATCAAGAGGAAAGAGTTTGTCTATGAAAGGTTAGGGTGTTCAAAG 88  
DB 51 TGCTGTGGGAGCAGTAGCAGGAAAGAGTTTGTCTAGAAAGACTCGGCTGCTTCAGTG 110  
QY 89 ATGTTTACCATGGACCGAGCTTCTCAACAGAGTTGGTAGGTTTACCTGCTCTCCAG 148  
DB 111 ATGACTCCCATGGTCAGGAATTTACGGAAGACCCCTCCATATATTGCTTGGTCTCCAA 170  
QY 149 AGAAGATAAACACTCGTTTCTGCTCTACACTATACAAATCCCAATGCCCTATCAGGAGA 208  
DB 171 AAGATGTCAACACCCGCTTCTCTCTATATATACTAATGAGAACCCAAACAACTTTCAGAG 230  
QY 209 TCAGTGGGTAAATTTCTTCAACTATCCAGCTCATATTTTGGAAACAGACAGATCACCC 268  
DB 231 T---TGCCGACAGATTCATCAAGCATCAGTGGCTCCAAATTTCAAAACAAATAGAAAACTC 287  
QY 269 GTATCAACATAGCTGGAT-----GGAAAAACAGATGGCAAAATGGCAGAGACATGTGCA 322  
DB 288 GCTTTTATTATTCATGATTCATAGACAGGAGAGAAACTGGCTGGCCATGTGTGCA 347  
QY 323 ATGCTGTGTACAGCTGGAGATATAAATTTGATTAATTTAGATTGGATCAACGGTTCCAC 382



348 AGAATCTGTTCAAGGTGGAAAGTGTGAACCTGTATCTGTGTGACTGGAAAGTGGCTCCC 407  
383 GG---GAATACATCCATGCTGTAAACAAATCTCGTGTCTGCTGCTGAGGTGGCTTATT 439  
408 GAACTGGATACACACAGCTCCGACAAATCAGGATCGTGGGACAGAAAGTGGCATATT 467  
440 TTATTGATGTTTCATGAAATAATTTGAATATTCCTCCCTTCTAAAGTGCACATTTGCC 499  
468 TTGTTGAATTTCTTCAGTCGGCTTCGGTTACTCACTTCCAACTGATGATGCTATGGCC 527  
500 ACAGCTTGGGAGCACACCTGGCTGGGGAGCTGGGTCAAGGATACAGG---CCTTGGAA 556  
528 ACAGCTTGGGAGCACACCTGGCTGGGGAGCTGGGTCAAGGATACAGG---CCTTGGAA 587  
557 GAATACTGGGTTGGACCCAGCTGGGCCATTTTCCACAACTCCAAAGGAAGTCAAGGC 616  
588 GCATCACAGGTTGGACCCAGCAGAACCTTGTCTTCAGGCGACACCTGAATAGTCCGAT 647  
617 TAGACCCCTCGGATGCCAATTTGTTGACGTTATTTACAAATGACATGCTGCTCTCT 676  
648 TGACCCCTCGGATGCCAATTTGTTGACGTTATTTACAAATGACATGCTGCTCTCT 707  
677 TTGAGCTTGGTGTGGAAACCATGTGCTGTGGGTCACTTTCAGCTTTTACCAAAATGGAG 736  
708 CCAATTTGGGTTTGGAAATGAGCCAGTCTGGGCCACCTAGATTCTTCCAAATGGAG 767  
737 GGAAGCACATGCCAGATGTGAAGCTTAATTTACACCTTTTACTGAAATTTTAACTTCAATG 796  
768 GAGTGAAATGCTCGATGATAAAGAAACATTT---CTCTCTCAGATTGTGACATAGACG 824  
797 CTTTACAAAGAAATGGCTCTCTTCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 856  
825 GAATCTGGGAGGACTCGAGCTTGTGGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 884  
857 ATGCTGAAGCAATCTTAAATCTGTGATGCTTATTTGCTTATCTTGTAGATCTTCAACAT 916  
885 ACATGATAGCATCGTCAACCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 944  
917 CTTTAAAGCAGGAATGCTCTCTTCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 976  
945 TCTTCTGCTCAAGAGTGTTCCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1004  
977 TTGCTGATGATTTTCACTTCAAAATATGAAGCTTAATGATCATTATTTTAAACA 1036  
1005 ATGCTGATGATATCTTGGGAAA---ACAAATGATGTGGGCCAGAAATTTATCTAGACA 1061  
1037 CAGGTGCTCTTCCCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1096  
1062 CTGGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1121  
1097 GCGAGTCACTCAAGAACTGTCTTCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1156  
1122 AAAGGTTAC---AGGACATACATGATTTCTTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1178  
1157 AGTTGCTCATCTGAGTGGAACTTTGAGCCAGGATGATTTACACAAATTTAATCGATG 1216  
1179 AGTATGAAATTTCAAGGCACTCTCAAAACAGATGATGATTTCAATTCATGATTTGACT 1238  
1217 CAGATGTTAACTGTTGAAACATTTACAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1276  
1239 CAGATGTTGATTTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1298  
1277 AAGATTCTCAGAAATAGTTGGGAGCAGAAATGTTGATTAATATACATCTGGGAAATATGGAT 1336  
1299 ACCCACTTTTACCTAGATGGGAGCATCCAGATTTATAGTGGAGC---AAATGTTGGAA 1355  
1337 ATAAATCTACCTTCTGTAGCAGCAATTTATGAGGACCTTAATTTCTCCAGAACTCGAAC 1396  
1356 AACAGTTCACTTCTGTAGTCCAGAAACCGTCAGGAGGAAAGTTCTGCTCACCTCCACAC 1415  
1397 CATGCTA 1403

Db 1416 CGTGTTA 1422

## RESULT 10

US-09-969-347-194  
; Sequence 194, Application US/09969347  
; Patent No. US20020115085A1  
; GENERAL INFORMATION:  
; APPLICANT: Ebner, Reinhard  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu  
; TITLE OF INVENTION: Sets  
; FILE REFERENCE: 689290-59  
; CURRENT APPLICATION NUMBER: US/09/969,347  
; CURRENT FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: US/60/237,598  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: US/60/237,604  
; PRIOR FILING DATE: 2000-10-03  
; NUMBER OF SEQ ID NOS: 318  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 194  
; LENGTH: 1450  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-969-347-194

Query Match 24.1%; Score 338; DB 10; Length 1450;  
Best Local Similarity 54.9%; Pred. No. 3.2e-87;  
Matches 778; Conservative 0; Mismatches 620; Indels 18; Gaps 5;

QY 1 ATGCTTGAATTTGGATTGTCATTCTTCTTTGGCACATCAAGAGAAAGAGTTT 60  
Db 1 ATGCTGCCCCCTTGGACCTCGGCCCTTCTCTGCTGCCACAGTCAGAGGAAAGAGGTC 60  
QY 61 TCTATGAAAGTTAGGTTGTTTCAAGATGGTTTACATGGACCCAGGACTTCTCAACA 120  
Db 61 TGCTACGACAACTTGGCTGCTTCTGATGAAAAAACCATGGGCGAGAACCTTCACGGA 120  
QY 121 GAGTTGGTAGGTTTACCTGCTCTCCAGAGAAATAAACAATCTGTTTCTGCTCTACAT 180  
Db 121 CCGTAAATTTACTTCCCTGCTCCCGGAGACATTCACACCCGCTTCTTCTGTATACA 180  
QY 181 ATACAAATCCCAATGCTATCAGAGATCAGTGGGTAAATTTCTCAACTATCCAGCC 240  
Db 181 AATGAAATCCAAACATCTCCAACTAATCAGTGGCAGGAAACAGACCATTTAGGCT 240  
QY 241 TCATATTTTGAACAGACAAAGATCACCCGTATCAACATAGCTGG-----ATGAAAAACA 294  
Db 241 TCAAACTTCCAACTGGACCGCAAGACACGCTTCATCATCCATGGCTTTTAGACAAGCG 300  
QY 295 GATGCAAAATGGCAGAGACATGTGCAATGTGTGCTACAGCTGGAGATATAAATTGC 354  
Db 301 GAGGACAGCTGGCCATCGGACATGTGCAAGAAATTTTGAAGTGGAGAGGTTGAACTGC 360  
QY 355 ATTAATTTAGATTGGATCAACGGTTTCAACGGGAA---TACATCCATGCTTAAACAATCTC 411  
Db 361 ATCTGTGTGACTGGAGGACGCGTCCCGGCGCATGTACACCCAGCCGCTGCAAAACATT 420  
QY 412 CGTGTGTGTGCTGAGGTGGCTTATTTTATTTGATGTTCTCATGAAAAAATTTGAATAT 471  
Db 421 CGGTTGTGTGGGCGGAGACAGCTTCTTAATCAAGACATGTGCGACGAGCTAGGTTAC 480  
QY 472 TCCCTTCTAAAGTGCATTTGATTGGCCACAGCTTGGGAGCACACCTGGCTGGGAGCT 531  
Db 481 AGCTTGGAGACGTGCTATCGGCCACAGCTGGGCGGACACAGCCCGCGGAGGCG 540  
QY 532 GGGTCAAGGATACACAGGCC---TTGGAAGAAATTAATCTGGTTGGACCCAGCTGGGCCATT 598  
Db 541 GCGAGGAGGTTGGGGGCGCGCTGGGCGAGGATCACAGGCTGGATCCAGCAGGCGCGTGC 600  
QY 589 TTCCACACACTCCAAAGGAGTCAAGCTAGACCCCTCGGATGCCAACTTTGTTGAGTT 648  
Db 601 TTCAGGATGAACCTGAGGAGGTTGCGTTGGATCCATCTGACCCGCTGTTTGGATGTG 660

649 ATTCAACAATGCGAGCTCGCATCCTCTTTGAGCTGGTGTGGAAACCAATTGATGCTTGT 708

661 ATTCAACAGAGATCTTCTCCCATAGTTCCTTCCTAGGTTTCGGAATGAGCCAAAGGTG 720

709 GGTCACTTGACTTTTACCCAAATGAGGAGGACACATGCCAGGATGTGAAGACTTAAT 768

721 GGCCATCTGGATTTCTTCCAAATGGAGGAAGGAATGCCCGATGTGAAGAA---AAAT 777

769 ACACCTTTACTCGAAATTAACCTCAATGCTTACAAAAGAAATGCGTTCCTTTCTTGAC 828

778 GTCCCTTTCAACCAATTACTGATTAATGATGGAATATGGGAAGGAATGTGTGGCTTTGTGCT 837

829 TGTAAACCATGCCGAAGTTTCAATTTTATGCTGAAGACATCTTAATCCTGATGCAATT 888

838 TGCAATCACCTAAGAACGTTCGAGTATTACTCAAGCAGCGTCTCAACCTGATGGCTTC 897

889 ATTGCTTAATCCTTTGATAGCTCAACATCTTTTAAAGCAGGAATATGCTCTTTTGTTC 948

898 CTGGGCTATCCTCTGTGCCTCTACGATGAGTTTCAGGAGAGTAAGTGTTCCTCTGTCCA 957

949 AAAGAAAGTGTGCCCAACAATGGGTCATTTTCTGATAGATTTTCACTTCAAAAATATGAAG 1008

958 GCTGAAGGATGCCCAAAATGGGCACTATGCTTGACCAATTTAAGGGGAAA---ACAAGT 1014

1009 ACTAATGGATCACTATTATTTTAAACACAGGTCCTCTTCCCCATTTGCCGTTGGAGG 1068

1015 GCTGTGGAAACAAGCCTTTTTCCTGAAACACAGGAGAGTGGTAACTTTACTAGTTGGAGA 1074

1069 CACAAATTTGCTGTTTAAACTCAGTGGGAAGCGAAGTCACCTCAAGGAACCTGCTTTCTTCGT 1128

1075 TATAAGTATCAGTCACACTTTCTGGAAAAGAGAAAGTGAATGGGTACATCGGATGCT 1134

1129 GTAGCGGGGCAATTGGGAAAATGGGGAGTTTGCATTTGTCAGTGGAAAACTTGAGCCA 1188

1135 TTGTATGGAAGTAATGAAACTCGAAACAATATGATTTTCAAAAGATCTCCTCAAAACA 1194

1189 GGCATGACTTACACAAAATTAATCGATGCGAGATTTAAGTTTGGAAAAATTACAAGTGT 1248

1195 GATGCACTCACAGTGTGCTATTGATGTGGATTTTAAATTTTGGAAAAATACAGAAAGTT 1254

1249 CAGTTTACTCGAAAAACAATTTGTTTGAAGATTTCTCAGAATAAGTTGGGAGCAAAATG 1308

1255 AAATTCCTCTCGAAACAACCGTGGGATAAATCTATCTGAGCCCAACTCGGGGCTTCCCAA 1314

1309 GTGATAAATACATCTGGGAATAATGGATATAAATCTACCTTCTGTAGCCAGACATTAAG 1368

1315 ATCACATGCGAAAGTGTGAAGATGGGACTGAGTATAATTTTGTAGCAGGCACTGTG 1374

1369 GGACCTAATATTCTCCAGAACCTGAAACCAATGCTAA 1404

1375 GAAGAAAACGTTCTGCAATCTCTTTACCTTGTTAA 1410

[UT 11  
 9-925-297-280  
 uence 280, Application US/03925297  
 tent No. US20020081659A1  
 NERAL INFORMATION:  
 PPLICANT: Rosen et al.  
 ITTLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ILE REFERENCE: FA105  
 RRENT APPLICATION NUMBER: US/09/925,297  
 RRENT FILING DATE: 2001-08-10  
 RIOR APPLICATION NUMBER: PCT/US00/05989  
 RIOR FILING DATE: 2000-03-08  
 RIOR APPLICATION NUMBER: 60/124,270  
 RIOR FILING DATE: 1999-03-12  
 UMBER OF SEQ ID NOS: 928  
 FTWARE: PatentIn Ver. 2.0  
 ? ID NO 280  
 .LENGTH: 1506  
 TYPE: DNA

; ORGANISM: Homo sapiens  
US-09-925-297-280

Query Match 24.1%; Score 338; DB 9; Length 1506;  
Best Local Similarity 54.9%; Pred. No. 3.2e-87;  
Matches 778; Conservative 0; Mismatches 620; Indels 18; Gaps 5;

Qy	1	ATGCTTGGAAATTGGATTGTGCAATCTCTGTTCTTTGGGCATCAAGAGGAAAAGAGTT	60
Db	27	ATGCTTGCOCCTTGGACCTCGGCCTTCTCCTGCTGCCACATGCAGAGGAAAAGAGTCT	86
Qy	61	TGCTATGAAGGTTAGGGTGTTCAAAGATGGTTTACCATTGGACACGAGCTTTCTCAACA	120
Db	87	TGCTACGGACAACCTTGGCTGCTTTCTGATGAAAACCATGGCAGGAACCCCTTCAGCGA	146
Qy	121	GAGTTGGTAGTTTACCCCTGGTCTCCAGAGAATAAACAACCTGGTTTCTGCTCTACACT	180
Db	147	CCTGTAAAAATTACTTCCCTGGTCCCGCAGCAATTGACACCGCTTTCTTCTGTACACA	206
Qy	181	ATACACAATCCCAATGCTATCAGGAGATCAGTGGCGGTTAATTCTTCAACTATCAAGCC	240
Db	207	AATGAATTCUAAACAATCCCAACTAATCACTGGCACGGAACCAAGACCACTTGGGCT	266
Qy	241	TCATATTTTGGAAACAGACAAGATCACCGGTATCAACATAGCTGG-----ATGGAACA	299
Db	267	TCAACTTCCAACCTGGACCGCAGACACGCTTCATCATCCATGGCTTCTTAGACAAGCG	326
Qy	295	GATGCCAATGGCAGAGACATGTGCAATGTGTTCTACAGCTGGAAGATATAAATTGC	354
Db	327	GAGGACAGCTGGCCATCGGACATGTGCAAGAAAATGTTTGAAGTGGAGAGGTGAACCTG	386
Qy	355	ATTAAATTTAGATTGGATCAACGGTTCAACGGGAA---TACATCATCTGCTGTAAACAATCTC	411
Db	387	ATCTGTGTGACTGGAGGCACGGTCCGGGCAATGTACACCAAGCCGTGCARAACATT	446
Qy	412	CGTGTGTGTGCTGAGGTGGCTTATTTTATGATGTTCTCATGAATAAATTTGAAAT	471
Db	447	CGGGTGTGTGGGCGGAGACAGCTTTCTTAATAACAAGCACTGTCGACGCAGCTGGGGTAC	506
Qy	472	TCCCTCTTAAGTGCATTTGATTGGCCACAGCTTGGGAGACACCTGGCTGGGAAAGCT	531
Db	507	AGCCTTGAGACGTGCGATGATCGGGCACAGCCTGGGCGGCACACGGCCGCGAGGCG	566
Qy	532	GGGTCAAGGATACAGGCC---TTGGAAGAAATAAATGGGTTGGACCCAGCTGGGCGCAATT	588
Db	567	GSCAGAGGCTGGGGGCGCGCGTGGCAGATCAACAGGCTGATCCAGCAGGCGCGTGC	626
Qy	589	TTCCACAACACTCCAAGGAGTCAAGCTAGACCCCTCGGATGCCAACTTTGTGTGACGTT	648
Db	627	TTCCAGGATGAACCTGAGGAGTTCGGTTGGATTCATCTGACGCGCTGTTGTGGATGTG	686
Qy	649	ATTCAACAATGAGCTGCATCTCTTTGAGCTTGGTTGGAAACCAATTGATGCTTGT	708
Db	687	ATTCAACAGATTCTTCTCCATAGTTCCTTCCCTAGGTTCCGGAATGAGCCAAAGGTG	746
Qy	709	GGTCATCTTGACTTTTACCCAAATGAGGGAAGCACATGCCAGGATGTGAAGACTTAATT	768
Db	747	GGCCATCTGGATTCTTTCCAAATGAGGMAAGGAAATGCCCGGATGTAAGAA---AAAT	803
Qy	769	ACACCTTTACTGAAATTTACTTCAATGCTTACAAAAGAAATGGCTTCCTCTTTGAC	828
Db	804	GTCTCTTCAACCATTACTGATATGATGAAATATGGGAAGGAATTTGGTGGCTTTGTGCT	863
Qy	829	TGTAAACCATGCCGAAGTTATCAATTTTATGCTGAAAGCATTTCTTAATCCCTGATGCAATT	888
Db	864	TGCAATCACCTAAGAAGCTTCGAGTATTACTCAAGCAGGCTCCTCAAACCTGATGGCTTC	923
Qy	889	ATTGCTTATCTTGTAGATCCTACATCTTTTAAAGCAGGAAATTTGCTTCTTTTGTTC	948
Db	924	CTGGGCTATCCCTGTGCTCTCTACGATGAGTTTCAGGAGTGAAGTGTTCCTCTGTCCA	983
Qy	949	AAAGAAGGTTGCCCAACAATGGGTCAATTTGCTGTGATAGATTTCACTTCAAAATAATGAAG	1008

984 GCTGAAGGATGCCCAAAATGGGGCACTATCTGACCAATTTAAGGGGAAA---ACAAGT 1040  
1009 ACTAATGGATCACATTAATTTTAAACACAGAGGTCCCTTTCCCAATTTGCCCGTTGGAGG 1068  
1041 GCTGTGGAACAAACCTTTTCCCTGAAACACAGGAGAGAGTGAACCTTTACTAGTTGGAGA 1100  
1069 CACAAATTTGCTGTAAACACTCAGTGAAGGAGTCACTCAAGGAACTGCTCTTTCTTGGT 1128  
1101 TATAAGGTATCAGTCACATTTCTGGAAGAAGAAAGTGAATGAGGTACATCAGGATTTCT 1160  
1129 GTAGGGGGCAATTTGGAAAACTGGGGAGTTTGCCATTTCTCAGTGGAAAACTTTCAGGCA 1188  
1161 TTGTATGGAAGTAATGAAACTCGAAACCAATATGAGATTTTCAAGAGATCCCTCAACCA 1220  
1189 GGCATGACTTTACACAAATTAATTCGATGAGATTTAACTGTTGGAAAACTTACAAAGTGT 1248  
1221 GATGCAAGTCACACGCTGTGCTATTTGATGTTGGAATTTTAAATTTGTTGGAATAATACAGAAAGTT 1280  
1249 CAGTTTCATCTGGAATAAACTTTGTTTGAAGATTTCTCAGAAATAGTTGGAGCAGAAATG 1308  
1281 AATTCCTCTGGAACAAACGTTGGATAAATCTATCTGAGCCCAAACTGGGGCTTCCCAA 1340  
1309 GTGATAAATACATCTGGAAATATGATATAAATCTACCTTCTGTAGCCAAACATTTATG 1368  
1341 ATCAGAGTGAAGAGTGGTGAAGATGGGAGTGAATATAATTTTGTAGCAGGCACACTGTG 1400  
1369 GGCCTTAATTTCTCCAGAACCTGAAACCATGCTAA 1404  
1401 GAAGAAACGCTTGTGCAATCTCTTTACCCTTTGTTAA 1436

T 12

-923-779-139  
ence 139, Application US/09923779  
ent No. US20020076721A1  
ERAL INFORMATION:  
PLICANT: Pyle, Ruth A.  
PLICANT: Xu, Jiangchun  
PLICANT: Kalos, Michael D.  
TLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
LE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER  
LE REFERENCE: 210121.553  
RENT APPLICATION NUMBER: US/09/923,779  
RENT FILING DATE: 2001-08-06  
MBER OF SEQ ID NOS: 155  
FTWARE: FastSeq for Windows Version 4.0  
ID NO 139  
ENGTH: 938  
YPE: DNA  
GANISM: Homo sapiens  
EATURE:  
AME/KEY: misc feature  
OCATION: 8, 16, 919, 929  
THER INFORMATION: n = A,T,C or G  
-923-779-139

ry Match 16.0%; Score 224; DB 9; Length 938;  
t Local Similarity 59.8%; Pred. No. 3.6e-54;  
ches 450; Conservative 1; Mismatches 286; Indels 15; Gaps 4;

29 TGTCTTTTGGCACATCAAGAGGAAAGTTTGTATGAAGAGTTAGGGTTTCAAAG 88  
104 TGTCTGTGGAGCAGTAGCAGGAAAGAGTTTGTACGAAAGACTCGGCTGCTCAGT 163  
89 ATGGTTTACCATGGACAGGACTTTCTCAACAGAGTTGGTAGGTTTACCCTGGTCTCAG 148  
164 ATGATCTCCCATGGTCAGGAATACGGAAGACCCCTCCATATATTGCTTGGTCTCAA 223  
149 AGAAGATAAAGACTGTTTCTGCTCTACATATACAAATCCCAATGCCCTATCAGAGA 208  
224 AAGATGTCACACCCCGCTTCCCTCTATATATTAATGAGAACCCCAACAACTTTCAAGAAG 283  
209 TCAGTSCGGTAAATTTCTTCAACTACTCAAGGCTCATATTTTGGACACAGACAGATCAACC 268

Db 284 T---TGCCGACAGATTCATCAAGCATCAGTGGCTCCAAATTTCAAAACAAATAGAAAACCTC 340  
Qy 269 GTATCAACATAGTGGAT-----GGAAACAGATGGCAAAATGGCAGAGACATGTGCA 322  
Db 341 GCTTTATTTATTCATGATTCATAGACAAGGGAAGAAAACCTGGCTGGCCAAATGTGTGCA 400  
Qy 323 ATGTGTTGTACAGCTGGGAAGATATAAATTTGCAATTAATTTAGATTTGGATCAACGGTTTCA 382  
Db 401 AGAATCTGTTCAAGTGGAAAGTGTGAATCTGTATCTGTGTGACCTGGAAGAGTGGCTCCC 460  
Qy 383 G---GGAATACATCCATCTGTAAACAATCTCCGTGTTGTTGGTGTGCTGAGTGGCTTATT 439  
Db 461 GAACTGGATACACACAAGCCTCGCAACATCAGGATCTGTGGAGCAGAAAGTGGCATATT 520  
Qy 440 TTATTGATGTTCTCATGAAAAAATTTGAATATTTCCCTTCTAAAGTGCACTTTGATTGGCC 499  
Db 521 TTGTGAATTTCTTCAGTCGGCGTTTCGGTTACTCACCTTCCAACTGATGTCTATTGGCC 580  
Qy 500 AAGCTTTGGAGCACACCTGGCTGGGAAGCTGGGTCAAGGATACAGG---CCTTGGAA 556  
Db 581 ACAGCCTGGGTGCCACGCTGTCTGGGAGGCTGGGAAGGAGAACCAATGGGACCATTTGAC 640  
Qy 557 GAATAACTGGTGTGGACCCAGCTGGGCCATTTTCCCAACACACTCCAAAGGAAGTCAAGC 616  
Db 641 GCATCACAGGTTTGAACCCAGCAGAACCTTGTCTTCAGGGCACCCTGAAATTAGTCCGAT 700  
Qy 617 TAGACCCCTCGGATCCCAACTTTTGTGACGTTATTTATACAAATGCAAGTGCATCTCTCT 676  
Db 701 TGGACCCAGCAGCATGCCAAATTTTGTGGATGTAATTTACACGGATGGTGGCCCCATAGTCC 760  
Qy 677 TTGAGCTTGTGTTGGAAACCATGATGCTTGTGTGCTCATCTTGACTTTTACCCAATGGAG 736  
Db 761 CCAATTTGGGGTTTGAATGAGCCAAATCTCGTGGGCCACCTAGATTTCTTTCCAAATGGAG 820  
Qy 737 GGAAGCAGATGCCAGGATGTGAAGACTTTAAT 768  
Db 821 GAGTGGAAATGCTGATGTAAGAAACATTT 852

RESULT 13

US-09-923-779-76/c  
; Sequence 76, Application US/09923779  
; Patent No. US20020076721A1  
; GENERAL INFORMATION:  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Kalos, Michael D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER  
; FILE REFERENCE: 210121.553  
; CURRENT APPLICATION NUMBER: US/09/923,779  
; CURRENT FILING DATE: 2001-08-06  
; NUMBER OF SEQ ID NOS: 155  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 76  
; LENGTH: 907  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 23, 104, 110, 763, 786, 792, 819, 822, 834, 836, 842, 843,  
; LOCATION: 892, 900  
; OTHER INFORMATION: n = A,T,C or G  
US-09-923-779-76

Query Match 14.9%; Score 209.2; DB 9; Length 907;  
Best Local Similarity 59.0%; Pred. No. 7.2e-50;  
Matches 434; Conservative 0; Mismatches 286; Indels 15; Gaps 4;  
Qy 46 AGAGAAAAGAAAGTTTGTATGAAGGTTAGGGTGTTCCTTCAAGAGTGTTCACCATGACC 105  
Db 771 AGGAAAAAAGTTTGTACGAAAGACTCGGCTGCTTCAAGTATGACTCCCATGGTCA 712

106 AGGACTTCTCAACAGAGTTGGTAGTTTACCTGGTCTCCAGAGAGATAAACAACCTGCT 165  
 711 GAATTAACGGGAAGACCCCTCCATATATGGCTTGGTCTCCAAAGAGTGTCAACACCGC 652  
 166 TTCTGCTCTACACTATACAAATCCCAATGCCATCAGAGAGTACAGTGGGTAAATCT 225  
 651 TTCTCTCTATATCTAATAGAACCCCAACAACATTTTCAAGAAGT---TGCCGAGATTCA 595  
 226 TCAACTATCCAAAGCTCATATTTTGGAAACAGACAGATCACCGTATCAACATAGCTGA 285  
 594 TCAGCATCAGTGGCTCCAAATTTCAAAACAAATAGAAAATCTGCTTTATTTATGAG 535  
 286 T-----CGAAAAACAGATGCGAAATGCGAGAGACATGTCGAATGTGTGTACAGTG 339  
 534 TTCTAGCAAGAGGAGAGAAATCTGCTGCGCAATGTGTGCAAGAATCTGTTCAAGGTG 475  
 340 GAAGATATAAATTCATTAATTTAGATTTGATCAACGGTTTCAG---GGAAATACATCAT 396  
 474 GAAGTGTGAATCTGTATCTGTGTGATGGAAGGTGGCTCCCGAACTGGATACACAA 415  
 397 GCTGTAAACAAATCCCGTGTGTGTGTGCTGAGTGGCTTTATTTATTTATGATGTTCTCATG 456  
 414 GCCTCGCAGAACATCAGGATCGTGGAGCAGAGTGGCATATTTTGTGAAATTTCTTCAG 355  
 457 AAAAAATTTGATATTTCCCTTTCTAAAGTGCATTTGATTTGGCCACACAGCTTGGAGCAC 516  
 354 TCGGGTTCGGTGTACTCACCTTCCAACTGGAAGGTGGCTCCCGAACTGGATGCGCCAC 295  
 517 CTGGTGGGGAAGCTGGGTCAAGATACACAG---CCTTGGAAAGAAATAACTGGGTGGAC 573  
 294 GCTGTGGGAGGCTGGAAGAGAACCAATGGGACCATTTGACGCAATCAGGGTTGGAC 235  
 574 CCAGTGGGCAATTTTCCACACACTCCTAAAGAGTACAGGTAGACCCCTGGATGCC 633  
 234 CCAGCAGAACCTTCTCTTCAAGGACACCTGAAATAGTCCGATTAGACCCCAAGCGATGCC 175  
 634 AACTTTGTGAGCTTATTATACAAATGCACTGCGATCCTCTTTGAGCTTGTGTGGA 693  
 174 AAATTTGTGATGTAAATTCACACGATGGTGGCCCAATAGTCCCAATTTGGGTTTGA 115  
 694 ACCATTGATGCTGTGTGATCTTTGATCTTTTACCCAAATGGAGGAGACATGCCAGGA 753  
 114 ATGACCAAGNCGTGGGCGCACTAGATTTCTTCCAAATGGAGGAGTGGAAATGCTCGGA 55  
 754 TGTGAAGACTTAAT 768  
 54 TGTAAAAAGAACATT 40

LT 14  
 9-923-779-94/c  
 quence 94, Application US/09923779  
 tent No. US20020076721A1  
 NERAL INFORMATION:  
 PPLICANT: Pyle, Ruth A.  
 PPLICANT: Xu, Jiangchun  
 PPLICANT: Kalos, Michael D.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER  
 FILE REFERENCE: 210121.553  
 CURRENT APPLICATION NUMBER: US/09/923,779  
 CURRENT FILING DATE: 2001-08-06  
 OFTWARE: FastSeq for Windows Version 4.0  
 Q ID NO 94  
 LENGTH: 879  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 106, 111, 649, 711, 779, 788, 808, 811, 812, 817, 820, 825,  
 LOCATION: 845, 851

OTHER INFORMATION: n = A,T,C or G  
 US-09-923-779-94

Query Match 14.5%; Score 204.2; DB 9; Length 879;  
 Best Local Similarity 58.8%; Pred. No. 28-48;  
 Matches 443; Conservative 0; Mismatches 294; Indels 16; Gaps 5;  
 QY 29 TGTCTTTGGCACATCAAGAGGAAAGAGTGTGTATGAAAGGTAGGGTGTTCACAA-A 87  
 DB 790 TCNTGCTGGGANCAGTAGCAGAAAGAGTGTGTACGAAAGAGACTCGGCTGCTTCAAGT 731  
 QY 88 GATGGTTTACCATGGAACAGGACTTTCTCAACAGAGTTGGTAGGTTTACCTGGTCTCCA 147  
 DB 730 GATGACTCCCATGGTGCAGNAATTAACGAAAGACCCCTCCATATATTGCTTGGTCTCCA 671  
 QY 148 GAGAAGATAAACAACCTCGTTTCTGCTCTACATATACACAATCCCAATGCCATCAGAG 207  
 DB 570 AAGATGTCACACACCGCTTCNTCTATATATATATGAGAACCCCAACACTTTCAGAA 611  
 QY 208 ATCAGTGGGTTTAAATCTTCAACTATCCAAAGCTCATATTTTGGAAACAGACAGATCACC 267  
 DB 610 GT---TGCCGCGAGATTCATCAAGCATCAGTGGCTCCAAATTTCAAAACCAATAGAAAAA 554  
 QY 268 CETHATCAACATACCTGAT-----GGAAAAACAGATGGCAATGCGAGAGACATGTGC 321  
 DB 553 CGCTTTATTTATTCATGATTCATAGCAAGGAGAGAAACTGCTGGCCAAATGTGTGC 494  
 QY 322 AATGTGTGTACAGCTGGAAGATATAAATTCATTAATTTAGATTGGATCAACGGTTTCA 381  
 DB 493 AAGAAATCTGTTCAAGGTGGAAGTGTGAACTGTATCTGTGTGACTGGAAGAGGTGGCTCC 434  
 QY 382 CG---GGAATACATCATCTGTAAACATCTCCGTGTGTGTGTGCTGAGGTGGCTTAT 438  
 DB 433 CGAACTGGATACACACACAGCTCGCGAAGAACATCAGATCTGGGAGCAGAGTGGCATAT 374  
 QY 439 TTTATTGATGTTCTCATGAAAAATTTGAATATTCCTCTCTTAAAGTGCATTTGATGGC 498  
 DB 373 TTTGTTGAATTTCTTCAGTGGGCTTGGTTTACTCACTTCCACGTCATGTCATTGGC 314  
 QY 499 CACAGCTTGGAGCACACCTGCTGGGAGTGGGTCAAGATACCCAG---CCTTGA 555  
 DB 313 CACAGCTGGTGGTCCACGCTGCTGGGAGGCTGGAAGAGAACCAATGGGACCAATTGA 254  
 QY 556 AGAATAACTGGGTGGACCCAGCTGGGCCATTTTCCACAAACATCTCCAAAGAGAGTCCAGG 615  
 DB 253 CGCATCACAGGTTGACCCAGCAGAACCTTCTTTCAGGGCACACCTGAATTAGTCCGA 194  
 QY 616 CTAGACCCCTCGGATGCCAACTTTGTGAGTTTATTCATAAAATGCGAGCTCGCATCTC 675  
 DB 193 TTGACCCCGAGCGATGCCAAATTTGTGGATGTAATTCACAGGATGGTGGCCCATAGTC 134  
 QY 676 TTTGAGCTTGGTGTGGAACCATTTGATGCTTGTGCTCATCTTGACTTTTACCCAATGGA 735  
 DB 133 CCCAAATTTGGGTTTGGATGANCCAATCTGTTGGCCACCTAGATTTCTTCCAAATGGA 74  
 QY 736 GGAAGACATGCGGATGTGAAGACTTAATT 768  
 DB 73 GAGTGGAAATGCTGATGTAAAGAGACATT 41

RESULT 15  
 US-09-923-779-22/c  
 ; Sequence 22, Application US/09923779  
 ; Patent No. US20020076721A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pyle, Ruth A.  
 ; APPLICANT: Xu, Jiangchun  
 ; APPLICANT: Kalos, Michael D.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER  
 ; FILE REFERENCE: 210121.553  
 ; CURRENT APPLICATION NUMBER: US/09/923,779  
 ; CURRENT FILING DATE: 2001-08-06

NUMBER OF SEQ ID NOS: 155  
 SOFTWARE: FastSeq for Windows Version 4.0  
 ID NO 22  
 LENGTH: 963  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: 26, 39, 40, 43, 44, 46, 54, 57, 948, 956  
 OTHER INFORMATION: n = A,T,C or G  
 -923-779-22

try Match 14.5%; Score 203; DB 9; Length 963;  
 Local Similarity 59.9%; Pred. No. 4.8e-48;  
 ches 452; Conservative 0; Mismatches 285; Indels 18; Gaps 6;  
 29 TGTCTCTTGGCACATCAAGAGGAAAGAGTTTGTCTATGAAAGTTTAGGGTGTTCCTCAAG 88  
 860 TGTCTGTGGGAGCAGTAGCAGGAAAGAGTTTGTCTACGAAAGACTCGCTGCTTCAGTG 801  
 89 ATGGTTTACCATGGACAGACTTTCTCAACAGAGTTGGTAGGT-TTACCCCTGGTCTCCA 147  
 800 ATGACTCCCCCATGGTCAGGAATTACGGAAGAGCCCTCCATATATTGCTTGGTCTCCA 741  
 148 GAGAAGATAAACAACCTGTTCTCTCTCTACACTATACACAATCCCAATGCCATACAGAG 207  
 740 AAGATGTCAACACCCGCTTCTCTCTATATCTAATGAGAACCCCAACACTTTCAAGAA 681  
 208 ATCAGTGGCGTTAATCTTCAACTATCCAGGCTCATATTTTGGAAACAGACAAGATCACC 267  
 680 GT---TGCCGACATTCATCAAGCATCAGTGGCTCCAATTTCAAAACAATAGAAAACT 624  
 268 CGTATCAACATAGCTGGAT-----GGAAACAGATGGCAATGGCAGAGAGACATGTGC 321  
 623 CGCTTTATTATTTCATGGATTCATAGCAAGGAGAGAAACTGGCTGGCCCAATGTGTGC 564  
 322 AATGTGTGTGTACAGCTGGAAGATATAAATTTGCATTAAATTTAGATTGGATCAACGGTTCA 381  
 563 AAGAACTGTTCAGGTGGAAAGTGTGAATGTATCTGTGTGACTGGAAGGTGGCTCC 504  
 382 CG---GGAATACATCCATGTGTAAACAATCTCCGTTGTGTGTGTGTGTGTGTGTGTGT 438  
 503 CGAATGGATACACAAAGCTCGCAGAACATCAGGATCGTGGGAGCAGAAAGTGCCATAT 444  
 439 TTTATTGATGTTCTCATGAAATAATTTGAATATCCCTTCTAAGTGCACCTTGTATTGGC 498  
 443 TTTGTGAATTTCTTCAGTGGGTTTCGGTTTACTCACCTTCCACGTGCATGTCAATTGGC 384  
 499 CACAGCTTGGGAGCACACCTGGCTGGGGAAGCTGGGTCAAGGATACCAAG---CCTTGA 555  
 383 CACAGCTGGGTCCCAAGCTGTGGGAGGCTGGAAGGAGAACCAATGGGACCATTTGA 324  
 556 AGAATACTGGTTGGACCCAGTGGGCCATTTTCCACAACACTCCAAAGAGTCAAG 615  
 323 CGCATCACAGGTTGGACCCAGCAGAACCTTCTTCAGGGGCACACCTGAAATTAGTCCGA 264  
 616 CTAGACCCCTCGGATGCCAATTTGTGTGACGTATTATACAAATGCAGCTCGCATCCTC 675  
 263 TTGGACCCAGCATGCCAAATTTGTGGATGTAAATTCACCGGATGGTCCCCCATAGTC 204  
 676 TTTGAGCTTGGTTGGAAACCA--TTGATGCTTGTGGTCACTTGAATTTTACCCAAATG 733  
 203 CCCAATTTGGGGTTTGGAAATGAGCCCAACTCCGTGGGCCACCTAGATTTCTTTCCAAATG 144  
 734 GAGGGAAGCAGATGCCAGGATGTGAAGACTTAATT 768  
 143 GAGGAGTGGAAATGCTTGGATGTAAAGAACATTT 109

Job completed: January 6, 2004, 11:20:35  
 Time: 486 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

cleic - nucleic search, using sw model

n: January 6, 2004, 08:32:01 ; Search time 5223 Seconds  
(without alignments)  
10996.959 Million cell updates/sec

US-10-038-517-1

ct score: 1404

ncc: 1 acgctgaattgtattgt.....agaacctgaaccatgctaa 1404

ng table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

hed: 2888711 seqs, 20454813386 residues

number of hits satisfying chosen parameters: 5777422

um DB seq length: 0

um DB seq length: 2000000000

processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

ase :

GenEmbl:\*

1: gb\_ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vi:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pi:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vi:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_pln:\*

35: em\_htg\_rod:\*

36: em\_htg\_nam:\*

37: em\_htg\_vrt:\*

38: em\_sy:\*

39: em\_hgo\_hum:\*

40: em\_hgo\_mus:\*

41: em\_hgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1404	100.0	1404	6	AX544122 Sequence
2	1402.4	99.9	2352	6	AX310490 Sequence
3	1402.4	99.9	2374	9	HSW804731
4	1399.4	99.7	1401	6	AX310491 Sequence
5	1115.8	75.5	1526	6	AX535973 Sequence
6	808	57.5	1540	6	AX405998 Sequence
7	386.8	27.5	1493	4	DOGPLIP
8	386.2	27.5	1481	6	AX337813 Sequence
9	386.2	27.5	1481	9	HUMLRP1P
10	386.2	27.5	1513	9	BC025784
11	376.2	26.8	1443	4	ECPL
12	373.4	26.6	1620	10	AF061274
13	370.2	26.4	1534	10	RNLAPRNA
14	345.8	24.6	2338	10	AF177402
15	344.2	24.5	1846	10	AF395870
16	344.2	24.5	1848	10	AF027293
17	344.2	24.5	2414	10	AF177403
18	342.8	24.4	2101	10	MCLPRTD2
19	342.6	24.4	1454	9	HUMLRP1P
20	342.6	24.4	1471	6	AX337796
21	342.6	24.4	1471	6	AX527618
22	342.6	24.4	1471	9	HUMLTGLIP
23	342.6	24.4	1515	9	BC014309
24	338	24.1	1450	6	AX337814
25	338	24.1	1450	9	HUMLRP2P
26	338	24.1	1512	9	BC005989
27	336.2	23.9	1374	10	MCTRIGLYP
28	330.2	23.5	1404	6	A24059
29	329.8	23.5	1702	10	CPTGH
30	328.2	23.4	1492	10	RATPANLI
31	326.4	23.2	1355	10	D88534
32	326.4	23.2	1355	10	D88535
33	324.6	23.1	1545	10	RATGLYLIPA
34	323.2	23.0	1555	10	MUSCTLL
35	301	21.4	1444	4	RABTRIL
36	298.2	21.2	1359	4	AF359386
37	297.8	21.2	1206	6	BD023826
38	276.4	19.7	1384	6	A24057
39	267.6	15.1	1484	5	AB070722
40	241	17.2	1437	5	BC043753
41	224	16.0	938	6	AX527612
42	209.2	14.9	907	6	AX527549
43	204.2	14.5	879	6	AX527567
44	203	14.5	963	6	AX527495
45	202.6	14.4	766	6	AX527546

ALIGNMENTS

RESULT 1	AX544122	1404 bp	DNA	linear	PAT 23-NOV-2002
AX544122	Sequence 1 from Patent WO02053753.				
LOCUS	AX544122				
DEFINITION	Sequence 1 from Patent WO02053753.				
ACCESSION	AX544122				
VERSION	AX544122.1	GI:25277625			
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	Yu, X. and Turner, C.A.				
AUTHORS	Human lipase and polynucleotides encoding the same				
TITLE	Patent: WO 02053753-A 1 11-JUL-2002,				
JOURNAL					

## Lexicon Genetics Incorporated (US)

JRES  
source  
1..1404  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
COUNT 413 a 274 c 303 g 414 t  
IN

ary Match 100.0%; Score 1404; DB 6; Length 1404;  
st Local Similarity 100.0%; Pred. No. 0;  
:ches 1404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCTTGGAAATTCGAAATTTGCAATCTCTTCTTGGCAATCAAGAGGAAAGAAAGTT 60  
1 ATGCTTGGAAATTCGAAATTTGCAATCTCTTCTTGGCAATCAAGAGGAAAGAAAGTT 60

61 TGCATGAAAGGTTAGGGTTTCAAGATGGTTTACCATGGACAGGACTTCTCAACA 120  
61 TGCATGAAAGGTTAGGGTTTCAAGATGGTTTACCATGGACAGGACTTCTCAACA 120

121 GAGTTGGTAGGTTTACCTGCTCCAGAGAGATAAACAACACTCGTTTCCTGCTCTACACT 180  
121 GAGTTGGTAGGTTTACCTGCTCCAGAGAGATAAACAACACTCGTTTCCTGCTCTACACT 180

181 ATACACAATCCCAATGCTATCAGAGATCAGTGGCGGTTAAATCTTCAACTATCCAAGCC 240  
181 ATACACAATCCCAATGCTATCAGAGATCAGTGGCGGTTAAATCTTCAACTATCCAAGCC 240

241 TCATATTTTGGAAACAGACAGATCACCGGTATCAACATAGCTGGATGGAAACAGATGCC 300  
241 TCATATTTTGGAAACAGACAGATCACCGGTATCAACATAGCTGGATGGAAACAGATGCC 300

301 AAATGGCAGAGACATGTGCAATGTCTGCTACAGCTGGAAGATATAAATGCAATTAAT 360  
301 AAATGGCAGAGACATGTGCAATGTCTGCTACAGCTGGAAGATATAAATGCAATTAAT 360

361 TTAGATTGGATCAAGGTTTCAAGGATATCATCTGTTAAACATCTCCGTGTTGTT 420  
361 TTAGATTGGATCAAGGTTTCAAGGATATCATCTGTTAAACATCTCCGTGTTGTT 420

421 GGTGCTGAGGTGGCTTATTTTATGATGTTCTCATGAAAAATTTTGAATATCCCTTCT 480  
421 GGTGCTGAGGTGGCTTATTTTATGATGTTCTCATGAAAAATTTTGAATATCCCTTCT 480

481 AAAGTGCATCTGATTGGCCACAGCTTGGAGACACACCTGGCTGGGAGAGCTGGTCAAGG 540  
481 AAAGTGCATCTGATTGGCCACAGCTTGGAGACACACCTGGCTGGGAGAGCTGGTCAAGG 540

541 ATACCAGGCTTGGAAAGATAAATCTGGGTGGACCCAGCTGGCCATTTTCCACAACACT 600  
541 ATACCAGGCTTGGAAAGATAAATCTGGGTGGACCCAGCTGGCCATTTTCCACAACACT 600

601 CCAAGAGAGTCAAGCTAGACCCCTCGATGCCAACTTTGTGACGTTATTCATAAAT 660  
601 CCAAGAGAGTCAAGCTAGACCCCTCGATGCCAACTTTGTGACGTTATTCATAAAT 660

661 GCAGCTCCGATCTCTTTGAGCTGGTTGGACCATTTGATGCTTGGTCTGCTCTGAC 720  
661 GCAGCTCCGATCTCTTTGAGCTGGTTGGACCATTTGATGCTTGGTCTGCTCTGAC 720

721 TTTTACCCAAATGGAGGAGACATGCCAGGATGTGAAGACTTAATTAACCTTTACTG 780  
721 TTTTACCCAAATGGAGGAGACATGCCAGGATGTGAAGACTTAATTAACCTTTACTG 780

781 AAATTTACTTCAATGCTTCAAAAAGAAATGCTTCTTGGATCTTAACCTATGCC 840  
781 AAATTTACTTCAATGCTTCAAAAAGAAATGCTTCTTGGATCTTAACCTATGCC 840

841 CGAAGTATCAATTTTATGCTGAAAGCAATCTTAATCCTGATGCAATTAATGCTTATCCT 900  
841 CGAAGTATCAATTTTATGCTGAAAGCAATCTTAATCCTGATGCAATTAATGCTTATCCT 900

QY 901 TGTAGATCTACACATCTTTTAAAGCAGGAAATTCCTTTCTTTTGTTCAAAAGAGTTGC 960  
DB 901 TGTAGATCTACACATCTTTTAAAGCAGGAAATTCCTTTCTTTTGTTCAAAAGAGTTGC 960

QY 961 CCAACATGGTTCATTTTCTGATAGATTTCACTTCAAAATATGAAGACTTAATGGATCA 1020  
DB 961 CCAACATGGTTCATTTTCTGATAGATTTCACTTCAAAATATGAAGACTTAATGGATCA 1020

QY 1021 CATTTATTTTAAACACAGGTCCTTTTCCCAATTTGCGCGTTGGAGGCACAAAATTTGTCT 1080  
DB 1021 CATTTATTTTAAACACAGGTCCTTTTCCCAATTTGCGCGTTGGAGGCACAAAATTTGTCT 1080

QY 1081 GTTAAACTCAGTGGAGCAAGTCACTCAAGAACTGTCTTCTTCGTGTAGGCGGGCA 1140  
DB 1081 GTTAAACTCAGTGGAGCAAGTCACTCAAGAACTGTCTTCTTCGTGTAGGCGGGCA 1140

QY 1141 ATTGGAAAACTGGGAGTTTGCCTATTTGCTAGTGAATACTTGAGCCAGGATGACTTAC 1200  
DB 1141 ATTGGAAAACTGGGAGTTTGCCTATTTGCTAGTGAATACTTGAGCCAGGATGACTTAC 1200

QY 1201 ACAAATTAATCGATCAGATGTTAAAGTGGAAAACTTACAGTGTTCATCTCTGG 1260  
DB 1201 ACAAATTAATCGATCAGATGTTAAAGTGGAAAACTTACAGTGTTCATCTCTGG 1260

QY 1261 ABAACACATTTGTTTCAAGATTTCTCAGATTAAGTGGAGCAGAAATGCTGATAATACA 1320  
DB 1261 ABAACACATTTGTTTCAAGATTTCTCAGATTAAGTGGAGCAGAAATGCTGATAATACA 1320

QY 1321 TCTGGAAATATGATATAAATCTACCTTCTGTAGCAAGACATTTATGGACCTAATAT 1380  
DB 1321 TCTGGAAATATGATATAAATCTACCTTCTGTAGCAAGACATTTATGGACCTAATAT 1380

QY 1381 CTCAGAACCTGAAACCATGCTAA 1404  
DB 1381 CTCAGAACCTGAAACCATGCTAA 1404

RESULT 2  
AR310490 2352 bp DNA linear PAT 12-JUN-2003

LOCUS AR310490 Sequence 1 from patent US 6558936.  
DEFINITION AR310490  
ACCESSION AR310490  
VERSION AR310490.1 GI:31703471  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
UNCLASIFIED.  
REFERENCE 1 (bases 1 to 2352)  
AUTHORS Khodadoust,M. and Kapeller-Libermann,R.  
TITLE Human lipase proteins, nucleic acids encoding them, and uses of  
JOURNAL Patent: US 6558936-A 1 06-MAY-2003  
FEATURES  
source location/qualifiers  
1..2352 /organism="unknown"  
BASE COUNT 705 a 431 c 491 g 722 t 3 others  
ORIGIN

Query Match 99.9%; Score 1402.4; DB 6; Length 2352;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1403; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTTGGAAATTTGGAATTTGTCATCTTCTTGGCAATCAAGAGGAAAGAAAGTT 60  
DB 125 ATGCTTGGAAATTTGGAATTTGTCATCTTCTTGGCAATCAAGAGGAAAGAAAGTT 184

QY 61 TGCATGAAAGGTTAGGGTTTCAAGATGGTTTACCATGGACAGGACTTCTCAACA 120  
DB 185 TGCATGAAAGGTTAGGGTTTCAAGATGGTTTACCATGGACAGGACTTCTCAACA 244

QY 121 GAGTTGGTAGGTTTACCTGCTCCAGAGAGATAAACAACACTCGTTTCTGCTCTACCT 180  
DB 245 GAGTTGGTAGGTTTACCTGCTCCAGAGAGATAAACAACACTCGTTTCTGCTCTACCT 304



```
181 ATACACAATCCCAATCCCTATCAGGAGATCAGTGGGTTAAATCTTCAACTATCCAGGCC 240
305 ATACACAATCCCAATCCCTATCAGGAGATCAGTGGGTTAAATCTTCAACTATCCAGGCC 364
241 TCATATTTTGGACACAGACAGATCACCCTATCAACATAGCTGGATGGAACAAACAGATGGC 300
365 TCATATTTTGGACACAGACAGATCACCCTATCAACATAGCTGGATGGAACAAACAGATGGC 424
301 AAATGGCAGAGACACATGTGCATGTGTGTCTACAGCTGGAAGATATAAATGCAATTAAT 360
425 AAATGGCAGAGACACATGTGCATGTGTGTCTACAGCTGGAAGATATAAATGCAATTAAT 484
361 TTAGATGGATCAACCGTTTCAAGGAATACATCCATGCTGTAAACAATCTCCGTGTGTT 420
485 TTAGATGGATCAACCGTTTCAAGGAATACATCCATGCTGTAAACAATCTCCGTGTGTT 544
421 GGTGCTGAGTGGCTTATTTTATTTGATGTTCATGAAAAAATTTGATATATCCCTTCT 480
545 GGTGCTGAGTGGCTTATTTTATTTGATGTTCATGAAAAAATTTGATATATCCCTTCT 604
481 AAATGGCAGAGACACATGTGCATGTGTGTCTACAGCTGGAAGATATAAATGCAATTAAT 540
605 AAATGGCAGAGACACATGTGCATGTGTGTCTACAGCTGGAAGATATAAATGCAATTAAT 664
541 ATACCAAGGCTTGGAGAGATTAATGGTGTGGACCCAGCTGGGCCATTTTCCCAACACT 600
665 ATACCAAGGCTTGGAGAGATTAATGGTGTGGACCCAGCTGGGCCATTTTCCCAACACT 724
601 CCAAGGAAGTCAAGGCTAGACCCCTCGGATGCCAATTTGTGACGTTATTCATCAACAAT 660
725 CCAAGGAAGTCAAGGCTAGACCCCTCGGATGCCAATTTGTGACGTTATTCATCAACAAT 784
661 GCAGCTGCATCTCTTTGAGCTGGTGTGGACCATGATGCTTGTGTCATCTTGAC 720
785 GCAGCTGCATCTCTTTGAGCTGGTGTGGACCATGATGCTTGTGTCATCTTGAC 844
721 TTTTACCCTTAAATGGAGGAGACATGCGAGATGTGAAGACTTAATTTACACCTTTACTG 780
845 TTTTACCCTTAAATGGAGGAGACATGCGAGATGTGAAGACTTAATTTACACCTTTACTG 904
781 AAATTTAACTCAATGCTTACAAAAAGAAATGGCTTCTTCTTTGACTGTAAACCATGCC 840
905 AAATTTAACTCAATGCTTACAAAAAGAAATGGCTTCTTCTTTGACTGTAAACCATGCC 964
841 CGAAGTTATCAATTTTATGCTGAAAGCACTTCTTATCTGATGATTTATGCTTATCCT 900
965 CGAAGTTATCAATTTTATGCTGAAAGCACTTCTTATCTGATGATTTATGCTTATCCT 1024
901 TGTAGATCCTACACATCTTTTAAAGCAGGAAATGCTTCTTCTTTGTTTCCAAAGAGGTTGC 960
1025 TGTAGATCCTACACATCTTTTAAAGCAGGAAATGCTTCTTCTTTGTTTCCAAAGAGGTTGC 1084
961 CCAACAAATGGGTCATTTGCTGATAGATTTCACTTCAAAAATATGAAGACTAATGGATCA 1020
1085 CCAACAAATGGGTCATTTGCTGATAGATTTCACTTCAAAAATATGAAGACTAATGGATCA 1144
1021 CATATATTTTAAACACAGGCTCTTCTTCCCATTTGCTTCTGAGGCAAAATTTGCT 1080
1145 CATATATTTTAAACACAGGCTCTTCTTCCCATTTGCTTCTGAGGCAAAATTTGCT 1204
1081 GTTAACTCAGTGGAGGAAAGTCACTCAAGAACTGCTTCTTCTGCTAGGCGGGGCA 1140
1205 GTTAACTCAGTGGAGGAAAGTCACTCAAGAACTGCTTCTTCTGCTAGGCGGGGCA 1264
1141 ATTGGAAACTGGGGAGTTTCCCATTTGCTGATGAAACTGAGGCGGCACTGACTTAC 1200
1265 ATTGGAAACTGGGGAGTTTCCCATTTGCTGATGAAACTGAGGCGGCACTGACTTAC 1324
1201 ACAAAATTAATCGATGCACTGTTAAACGTTGGAAACATTAACAAGTGTTCAGTTTCATCGG 1260
1325 ACAAAATTAATCGATGCACTGTTAAACGTTGGAAACATTAACAAGTGTTCAGTTTCATCGG 1384
```

```
1261 AAAAAACATTTTGTGAAGATTTCTCAGAAATAGTTGGAGCAGAAATGGTGATAAATACA 1320
1385 AAAAAACATTTTGTGAAGATTTCTCAGAAATAGTTGGAGCAGAAATGGTGATAAATACA 1444
1321 TCTGGAAATATGGATATTAATCTACCTTCTGTAGCCAAAGACATTTGGGACCTTAATTT 1380
1445 TCTGGAAATATGGATATTAATCTACCTTCTGTAGCCAAAGACATTTGGGACCTTAATTT 1504
1381 CTCCAGAACCTGAAACCATGCTAA 1404
1505 CTCCAGAACCTGAAACCATGCTAA 1528

RESULT 3
HSM804731
LOCUS Homo sapiens mRNA; cDNA DKFZp313P1022 (from clone DKFZp313P1022).
DEFINITION Homo sapiens mRNA; cDNA DKFZp313P1022 (from clone DKFZp313P1022).
ACCESSION AL833418
VERSION AL833418.1 GI:21734059
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2374)
AUTHORS Bahr,A., Lauber,J., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp313P1022) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
Location/Qualifiers
1..2374
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp313P1022"
/cisue_type="human adipose"
/clone_lib="313 (synonym: hlcc2). Vector pTriplex2; host
DH10B; sites SfiI + SfiIIB"
/dev_stage="adult"
polya_site 2355
BASE COUNT 718 a 425 c 498 g 733 t
ORIGIN
Query Match 99.9%; Score 1402.4; DB 9; Length 2374;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1403; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCTTGAATTTGGATTTGCAATCTTCTTGTGGACATCAAGAGAAAGAGTT 60
DB 151 ATGCTTGAATTTGGATTTGCAATCTTCTTGTGGACATCAAGAGAAAGAGTT 210
QY 61 TGCTATGAAGAGTTAGGTTTCAAGAGATGGTTTACCATGGACAGGACTTTCTCAACA 120
DB 211 TGCTATGAAGAGTTAGGTTTCAAGAGATGGTTTACCATGGACAGGACTTTCTCAACA 270
QY 121 GAGTTGGTAGGTTTACCTGCTCCAGAGAGATTAACACTGTTTCTGCTCTACACT 180
DB 271 GAGTTGGTAGGTTTACCTGCTCCAGAGAGATTAACACTGTTTCTGCTCTACACT 330
QY 181 ATACACAATCCCAATCCCTATCAGGAGATCAGTGGGTTAAATCTTCAACTATCCAGGCC 240
DB 331 ATACACAATCCCAATCCCTATCAGGAGATCAGTGGGTTAAATCTTCAACTATCCAGGCC 390
```



601 CCAGAGGAGTCAGCTAGACCCCTCGGATGCCAATCTTGTGACGTTATTCATCAAAAT 660  
601 CCAGAGGAGTCAGCTAGACCCCTCGGATGCCAATCTTGTGACGTTATTCATCAAAAT 660  
661 GCAGCTCGCATCTCTTTGAGCTTGTGTGGAAACCAATGATGCTTGTGTCATCTTGAC 720  
661 GCAGCTCGCATCTCTTTGAGCTTGTGTGGAAACCAATGATGCTTGTGTCATCTTGAC 720  
721 TTTTACCCAAATGGAGGGAAGCAGATGCCAGGATGTGAAGACTTAATTAACACTTACTG 780  
721 TTTTACCCAAATGGAGGGAAGCAGATGCCAGGATGTGAAGACTTAATTAACACTTACTG 780  
781 AAATTAATCTCAATGCTTACAAAAGAAATGCTTCTTCTTTGATGCTTAACCAATGCC 840  
781 AAATTAATCTCAATGCTTACAAAAGAAATGCTTCTTCTTTGATGCTTAACCAATGCC 840  
841 CGAGCTTATCAATTTATGCTGAAGCAATCTTAATCCTGATGCAATTTATGCTTATCCT 900  
841 CGAGCTTATCAATTTATGCTGAAGCAATCTTAATCCTGATGCAATTTATGCTTATCCT 900  
901 TGTAGATCCTACATCTTTTAAAGCAGGAAATGCTTCTTTTGTTCAAAAGAGGTTC 960  
901 TGTAGATCCTACATCTTTTAAAGCAGGAAATGCTTCTTTTGTTCAAAAGAGGTTC 960  
961 CCAACAATGGGTCATTTGCTGATAGATTCACATTCAAAATATGAAGACTTAATGATCA 1020  
961 CCAACAATGGGTCATTTGCTGATAGATTTCACTTCAAAAATATGAAGACTTAATGATCA 1020  
1021 CATTATTTTAAACACAGGCTTCTTCCCATTTGCGGCTTGGAGGCAAAATGTCT 1080  
1021 CATTATTTTAAACACAGGCTTCTTCCCATTTGCGGCTTGGAGGCAAAATGTCT 1080  
1081 GTTAACTCAGTGGAGGAAAGTCACTCAAGGAACTGCTTCTTCTGTTAGGCGGGCA 1140  
1081 GTTAACTCAGTGGAGGAAAGTCACTCAAGGAACTGCTTCTTCTGTTAGGCGGGCA 1140  
1141 ATTGGAAACTGGGGAGTTGGCAATTCAGTGGAAATCTTGAGCAGGCACTGATTC 1200  
1141 ATTGGAAACTGGGGAGTTGGCAATTCAGTGGAAATCTTGAGCAGGCACTGATTC 1200  
1201 ACAAAATTAATCGATGCAATGTAACCTTGGAAACATTAAGTGTTCATCTGG 1260  
1201 ACAAAATTAATCGATGCAATGTAACCTTGGAAACATTAAGTGTTCATCTGG 1260  
1261 AAAAAACATTTGTTGAAGTCTCAGATTAAGTGGAGCAGAAATGTTGATTAATACA 1320  
1261 AAAAAACATTTGTTGAAGTCTCAGATTAAGTGGAGCAGAAATGTTGATTAATACA 1320  
1321 TCTGGAAATATGGATATAAATCTACCTTCTGTAGCCAAAGACATTTAGGACCTTAAT 1380  
1321 TCTGGAAATATGGATATAAATCTACCTTCTGTAGCCAAAGACATTTAGGACCTTAAT 1380  
1381 CTCAGAACCTGAAACCATGC 1401  
1381 CTCAGAACCTGAAACCATGC 1401

T 5  
973  
AX535973 1526 bp DNA linear PAT 22-NOV-2002  
Sequence 12 from Patent WO0229036.  
ION  
AX535973  
ON AX535973.1 GI:25262503  
RDS  
E Homo sapiens (human)  
ANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Harland, L., Arvizu, C., Das, D., Griffin, J.A., Baughn, M.R., Ding, L.,  
Wallis, N.K., Yao, M.G., Lu, Y., Elliott, V.S., Thangavelu, K.,  
Pankumar, J., Lal, P.G. and Tribouley, C.M.  
Lipid metabolism enzymes

JOURNAL Patent: WO 0229036-A 12 11-APR-2002;  
Incyte Genomics, Inc. (US)  
FEATURES Location/Qualifiers  
source 1..1526  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/notes="Incyte ID No: 7482937CB1"  
BASE COUNT 440 a 303 c 325 g 458 t  
ORIGIN  
Query Match 79.58; Score 1115.8; DB 6; Length 1526;  
Best Local Similarity 94.33; Pred. No. 3.7e-260; Indels 71; Gaps 1;  
Matches 1198; Conservative 0; Mismatches 2;  
QY 205 GAGATCAGTGGGTAAATCTTCAAACTATCCAAAGCTCATATTTTGGAAACAGACAGATC 264  
DB 95 GAGATCAGTGGGTAAATCTTCAACTATCCAAAGCTCATATTTTGGAAACAGACAGATC 154  
QY 265 ACCGATATCAACATAGCTGGATGGAAAACAGATGGCAATGGCAGAGACATGTCAAT 324  
DB 155 ACCGATATCAACATAGCTGGATGGAAAACAGATGGCAATGGCAGAGACATGTCAAT 214  
QY 325 ----- 324  
DB 215 GTATGACATGAATGAAGCTCCTTTTACACTAGCATGGAGCTTTATGTTTAACATGAATG 274  
QY 325 -----GTGTTCTACAGCTGGAGATATAAATTGCATTAATTTAGATTTGATCA 373  
DB 275 TACTTTGCAAGGTGTTCTACAGCTGGAGATATAAATTGCATTAATTTAGATTTGATCA 334  
QY 374 ACGTTTCAACGGGATACATCCATGCTGTAAACAATCTCGTGTGTTGGTGTGCTGAGGTGG 433  
DB 335 ACGTTTCAACGGGATACATCCATGCTGTAAACAATCTCGTGTGTTGGTGTGCTGAGGTGG 394  
QY 434 CTTATTTTATGATGTTCTCATGAATAAATTTGAATATTTCCCTTTCTAAAGTGCATTTGA 493  
DB 395 CTTATTTTATGATGTTCTCATGAATAAATTTGAATATTTCCCTTTCTAAAGTGCATTTGA 454  
QY 494 TTGGCCACAGCTTGGGAGCACACCTGGCTGGGAGCTGGTCAAGGATACCAAGGCTTG 553  
DB 455 TTGGCCACAGCTTGGGAGCACACCTGGCTGGGAGCTGGTCAAGGATACCAAGGCTTG 514  
QY 554 GAAGAATAAATCGGCTTGGAGCCAGCTGGGCCATTTTCCACAACACTCCAAAGGAAATCA 613  
DB 515 GAAGAATAAATCGGCTTGGAGCCAGCTGGGCCATTTTCCACAACACTCCAAAGGAAATCA 574  
QY 614 GGCTAGACCCCTCGGATGCCAACTTTGTTGACGTTATTCATAAATGAGCTCGCATCC 673  
DB 575 GGCTAGACCCCTCGGATGCCAACTTTGTTGACGTTATTCATAAATGAGCTCGCATCC 634  
QY 674 TCTTTGAGCTTGGTGTGGAAACCAATGATGCTTGTGTTGATCTTTTACCCAAATG 733  
DB 635 TCTTTGAGCTTGGTGTGGAAACCAATGATGCTTGTGTTGATCTTTTACCCAAATG 694  
QY 734 GAGGGAAGCAGATGCCAGGATGTGAAGACTTAATTAACCTTTTACGAAATTTAATCTCA 793  
DB 695 GAGGGAAGCAGATGCCAGGATGTGAAGACTTAATTAACCTTTTACGAAATTTAATCTCA 754  
QY 794 ATGCTTACAAAAGAAATGCTTCTTCTTGTGCTGTAACCAATGCCCGAAGTTATCAAT 853  
DB 755 ATGCTTACAAAAGAAATGCTTCTTCTTGTGCTGTAACCAATGCCCGAAGTTATCAAT 814  
QY 854 TTTATGCTGAAGCAATTTTAACTCCTGATGCAATTTATTTGCTTATCTTTGATGATCTTACA 913  
DB 815 TTTATGCTGAAGCAATTTTAACTCCTGATGCAATTTATTTGCTTATCTTTGATGATCTTACA 874  
QY 914 CATCTTTTAAAGCAGGAAATGCTTCTTCTTGTGCTGTAACCAATGCCCGAAGTTATCAAT 973  
DB 875 CATCTTTTAAAGCAGGAAATGCTTCTTCTTGTGCTGTAACCAATGCCCGAAGTTATCAAT 934  
QY 974 ATTTTGTGATAGATTTTCACTTCAAAAATATGAAGACTTAATGATGATCAATTTTATTA 1033

935 ATTTTGTGATGATGATTTCACTTCAAAAATATGAGACTAATGATCACATTTATTTTAA 994  
 1034 ACACAGGTCCTTTTCCCATTTGCGCGTTGGAGGCACAAATTTCTGTTAACTCAGTG 1093  
 995 ACACAGGTCCTTTTCCCATTTGCGCGTTGGAGGCACAAATTTCTGTTAACTCAGTG 1054  
 1094 GAACGGAAGTCACTCAAGGAACCTGTTCTTCTGTTAGCGGGGCAATTTGGAAAACTG 1153  
 1055 GAACGGAAGTCACTCAAGGAACCTGTTCTTCTGTTAGCGGGGCAATTTGGAAAACTG 1114  
 1154 GGGAGTTGGCATTTGAGTGGAAACCTTGGAGGCACAAATTTCTGTTAACTCAGTG 1213  
 1115 GGGAGTTGGCATTTGAGTGGAAACCTTGGAGGCACAAATTTCTGTTAACTCAGTG 1174  
 1214 ATGAGATGTTTAACTGTTGGAACCAATTAACAAGTCTTCACTGTTGTTGTTGTTGTT 1273  
 1175 ATGAGATGTTTAACTGTTGGAACCAATTAACAAGTCTTCACTGTTGTTGTTGTTGTT 1234  
 1274 TTGAGATTTCTCAGAAATGTTGGAGGCACAAATTTCTGTTAACTCAGTGTTGTTGTT 1333  
 1235 TTGAGATTTCTCAGAAATGTTGGAGGCACAAATTTCTGTTAACTCAGTGTTGTTGTT 1294  
 1334 GATATAATCTACCTTCTGTTAGCAAGACATTTATGGAACCTTATTTCTCCAGAACCTGA 1393  
 1295 GATATAATCTACCTTCTGTTAGCAAGACATTTATGGAACCTTATTTCTCCAGAACCTGA 1354  
 1394 AACCATGCTAA 1404  
 1355 AACCATGCTAA 1365

5 AX405998 1540 bp DNA linear PAT 14-JUN-2002  
 NITION Sequence 413 from Patent WO0222660.  
 AX405998  
 AX405998.1 GI:21439435  
 DRDS Homo sapiens (human)  
 ZANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 Tang, X.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,  
 Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T.  
 Novel nucleic acids and polypeptides  
 Patent: WO 0222660-A 413 21-MAR-2002;  
 HYSEQ, INC. (US)  
 JRES Location/Qualifiers  
 1. 1540  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 86...916  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD35105.1"  
 /db\_xref="GI:21439435"  
 /translation="MLGIVAVFLPFGTSRGKVCVERLGCDFKGLPWTRTFSTELVG  
 LPSPKINTRFLYTHNPNAYQEIYSAVNSTIOASVEGDKITRINLAKGKTIDGW  
 QRDWNVLLQEDINCINLWINGSREYIHAVNNRVGAEVAFIDVLMKFEYSFS  
 KVHLHSHLAGHLAGISPIGLRITGLDPAGFPFHTPKREVDPSDFDVVH  
 TNAARILFELGVGTIDACGHLDFYPNGKHPMGCEDLTLLKFNFNAYKVENLSQA  
 "

CDS  
 477 a 276 c 324 g 463 t  
 COUNT  
 IN  
 ery Match 57.5%; Score 808; DB 6; Length 1540;  
 st Local Similarity 100.0%; Pred. No. 28-185;  
 tches 808; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 ATGCTTGGAAATTTGGATTTGCTTCTGTTTGGACATCAAGAGAAAAGAGTT 60  
 |||||

Db 86 ATGCTTGGAAATTTGGATTTGCTTCTGTTTCTTTGGACATCAAGAGAAAAGAGTT 145  
 QY 61 TGTATGAAAGTTTGGGTTTCAAGATGTTTACCATGTTTACCAGGACTTTTCTCAACA 120  
 Db 146 TGTATGAAAGTTTGGGTTTCAAGATGTTTACCATGTTTACCAGGACTTTTCTCAACA 205  
 QY 121 GAGTTCGTAGTTTACCTTGGTCTCCAGAGAATAAACAACCTCGTTTCTGCTTACACT 180  
 Db 206 GAGTTCGTAGTTTACCTTGGTCTCCAGAGAATAAACAACCTCGTTTCTGCTTACACT 265  
 QY 181 ATACACAATCCCAATGCTTATCAGGAGATCAGTGGGTAAATTTCAACTATCCAGGCC 240  
 Db 266 ATACACAATCCCAATGCTTATCAGGAGATCAGTGGGTAAATTTCAACTATCCAGGCC 325  
 QY 241 TCATATTTTGGAAACGACAGATCACCGGTATCAACATAGCTGGATGGAAACAGATGGC 300  
 Db 326 TCATATTTTGGAAACGACAGATCACCGGTATCAACATAGCTGGATGGAAACAGATGGC 385  
 QY 301 AAATGGCAGAGACATGTCATGCTGTGTACAGCTGGAAGATATAAATTTGCATTAAT 360  
 Db 386 AAATGGCAGAGACATGTCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 445  
 QY 361 TTAGATTTGGATCAACGGTTTCAAGGGAATACATCCATGCTGTAAACAATCTCCGTGTTGT 420  
 Db 446 TTAGATTTGGATCAACGGTTTCAAGGGAATACATCCATGCTGTAAACAATCTCCGTGTTGT 505  
 QY 421 GGTGCTGAGTGGCTTATTTTATGATGTTCTCATGAAAAAATTTGAATATTCCTCTCT 480  
 Db 506 GGTGCTGAGTGGCTTATTTTATGATGTTCTCATGAAAAAATTTGAATATTCCTCTCT 565  
 QY 481 AAAGTGCATTTGATGGCCACAGCTTGGAGACACACCTTGGCTGGGGAAGCTGGGTCAAGG 540  
 Db 566 AAAGTGCATTTGATGGCCACAGCTTGGAGACACACCTTGGCTGGGGAAGCTGGGTCAAGG 625  
 QY 541 ATACCAGGCTTGGAAAGATAAATCTGGTGTGGACCCAGCTGGGCCAATTTTCCACACACT 600  
 Db 626 ATACCAGGCTTGGAAAGATAAATCTGGTGTGGACCCAGCTGGGCCAATTTTCCACACACT 685  
 QY 601 CCAAGGAAGTCAAGGCTGAGCCCTCGGATGCGACACCTTGTGTGAGCTTATTTCATACAAT 660  
 Db 686 CCAAGGAAGTCAAGGCTGAGCCCTCGGATGCGACACCTTGTGTGAGCTTATTTCATACAAT 745  
 QY 661 GCAGCTCGCATCTCTTTGAGCTTGGTGTGGAAACCAATGATGCTGTGGTGTGGCTTTCAC 720  
 Db 746 GCAGCTCGCATCTCTTTGAGCTTGGTGTGGAAACCAATGATGCTGTGGTGTGGCTTTCAC 805  
 QY 721 TTTTACCAATGGAGGAGACACATGCGAGGATGGAAGCTTAAATACACCTTACTG 780  
 Db 806 TTTTACCAATGGAGGAGACACATGCGAGGATGGAAGCTTAAATACACCTTACTG 865  
 QY 781 AAATTTAACTTCAATGCTTACAAAAAG 808  
 Db 866 AAATTTAACTTCAATGCTTACAAAAAG 893  
 RESULT 7  
 DOGFLIP  
 LOCUS  
 DEFINITION Canine lipase mRNA, complete cds.  
 ACCESSION M35302  
 VERSION M35302.1 GI:164047  
 KEYWORDS lipase.  
 SOURCE Canis familiaris (dog)  
 ORGANISM Canis familiaris  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 1 (bases 1 to 1493)  
 REFERENCES  
 AUTHORS Kerflec, B., LaPorte, K.S., Puigserver, A. and Scheele, G.  
 TITLE Primary structures of canine pancreatic lipase and phospholipase A2  
 JOURNAL messenger RNAs  
 MEDLINE Pancreas 1 (5), 430-437 (1986)  
 PUBMED 87175472  
 3562437

INT Original source text: Canine pancreas, cDNA to mRNA.  
 Draft entry and computer-readable [or printed] sequence for [1]  
 kindly submitted by G.Scheele, 06-MAR-1990.

RES Location/Qualifiers  
 source 1..1493  
 /organism="Canis familiaris"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9615"  
 <1..1493  
 /product="lipase mRNA"  
 35..1438  
 /note="lipase precursor"  
 /codon\_start=1  
 /protein\_id="AAA30885.1"  
 /db\_xref="GI:164048"  
 /translation="MVSIVTIALFLGLRAKAEVCEYQIGCFSDAEPWAGTAIRPLKV  
 LPMSPERIGTRFLLYNKNPNFQTLPSDPTIEASNFQDKTRFTTHGFKNGEE  
 NWLLDMCKMFKVEEVNCTICVMKGSOTSXTOAANNRVVGAOVAOMLSLANSYV  
 SPQVOLIGHSLGAHVAGAGSRTPLGRITGLDPEASFOGTREVRPLDPTDADFDV  
 VIHTDAPIPLFGTGSQMGHLDFPFNGEEMEGCKNALSIQVINDGLWGETRDF  
 VACNHLRSKYXSESLNPDGFASYPASIRAFESNKCFPCPDQCGPOMGHYADFPV  
 KTSDETQKFLANTGSSNPFARWRYGVSTILSGKRATGQAKVALFGSKGNTHQFNIFKG  
 ILKPGSTHNEFDKLDVGTIEKVKFLMNNVNVNPTPEKVGAAKITVQGBEKTVHSF  
 CSBSTVREDVLLTLTPC"

sig\_peptide 35..85  
 /note="lipase signal peptide"

mat\_peptide 86..1435  
 /product="lipase"

COUNT 410 a 377 c 378 g 328 t

try Match 27.5%; Score 386.8; DB 4; Length 1493;  
 it Local Similarity 57.2%; Pred. No. 3.8e-83;  
 :ches 807; Conservative 0; Mismatches 587; Indels 18; Gaps 5;

1 ATGCTGGAAATTTGGATTGTTCATCTCTTTGTTGTCACATCAAGAGGAAAGAAAGTT 60  
 35 ATGGTAGCATCTGGCAATCGCGCTGTTCTGCTGGGAGCAGCCAAAGCAAGAAAGTT 94  
 61 TGCTATCAAGAGTTAGGTTTCAAGATGGTTTACCATGACGACGAGCAATCTTCAACA 120  
 95 TGCTATCAAGCAATCGGATGCTTTCTGACGGAGGACCTGGCGCGGAGACCAATCAGG 154  
 121 GAGTTGTAGGTTTACCTGGTCTCCAGAGAAATAACACACTCGTTTCTGCTCTACACT 180  
 155 CCTCTGAAGTTCTCCCTGGAGCCCGGAGAGATCGGCACCGCTTCTGCTCTACACC 214  
 181 ATACACAATCCCAATGCTATCAGGAGATCAGTGGGTTAAATCTTCAACTATCCAGGCC 240  
 215 AACAGAAACCAACAACTTTCAAACTCTCTTCCCTCTGATCCATCAACGATTGAGGCA 274  
 241 TCATATTTTGGACAGACAGATCACCCGATCAACATAGCTGGATGGAATA-----CA 294  
 275 TCAATTTTCAACAGAACAGAGACCCGGTTCCACATTCATGGCTTCTATAAAGAGGA 334  
 295 GATGGCAATGGCAGAGAGACATGTGCAATGTGTTGCTACAGCTGGAAGATATAAATTC 354  
 335 GAAGAGAACTGGCTGCTGGATATGTGCAAGAAATGTTCAAGGTTGAGGAGTGAATGTC 394  
 355 ATTAATTTAGATTGGATCAACGGTTC---ACGGGAATACATCCATGCTTAAACATCTC 411  
 395 ATCTGCGTGACTGGAAGAAAGTTTCCAAACCTCATATACACAGCTCCCAACACGTT 454  
 412 CGTGTGTTGCTGCTGAGGTGGCTTATTTTATTTGATGTTCTCATGAAATAATTTGAATAT 471  
 455 CGGGTGTGGCGCCAGGTGGCCAGATGCTCAGCATGCTCTCGGCAAACTACAGTAC 514  
 472 TCCCTTCTTAAGTGCACTTGAATTGGCCACAGCTTGGGAGCACACCTGCTGGGAAGCT 531  
 515 TCACCTTCCCAAGTCCAGCTCATCGGCCACAGCTAGGAGCCACCGTGGCTGGGAGGCA 574  
 532 GGGTCAAGGATACAGGCCCTTGGAAAGAAATAACTGGGTGGACCCAGCTGGGCCATTTTC 591

Db 575 GGGAGCAGGACTCCAGGTCTGGGAGGATTACAGGGTTGGATCTCTGTAGAACAGTTTC 634  
 Qy 592 CACAACTCCCAAGAGAGTCAAGCTAGACCCCTCGGATGCAAACTTTGTTGAGTTATT 651  
 Db 635 CAGGCACTCTGAGAGAGTTGCACTTGCATCCACTGATGCTGATTTGTTGATGATT 694  
 Qy 652 CATACAAATGCACTCGCATCTCTTTGAGCTTGGTGTGTGGAACCATTTGATGCTTGGT 711  
 Db 695 CACACAGATGCACTCCCTGATCCATTCCTGGGTTTGGAAACAAGCAACAGATGGT 754  
 Qy 712 CATCTTGACTTTTACCCAAATGGAGGAGACACATGCCAGATGTGAAGACTTAATACA 771  
 Db 755 CACCTTGACTTCTCCCAATGGTGGAGAGAAATGCCAGATGCAAGAA---GAATGCC 811  
 Qy 772 CTTTACTGAATTTAATCTCAATCTTACAAAAGAAATGGCTTCCTTTTGGACTGT 831  
 Db 812 CTGTACAGATCGTGAACCTAGATGGCATTTGGGAAGAACTCGGAGCTTCGTGGCTTGC 871  
 Qy 832 ACCATGCCCGAGTTATCAATTTTATGCTGAAGCAATCTTAATCTCTGATGCAATTTATT 891  
 Db 872 AATCACTGAGAAGTTTACAAGTATTACTCAGAGAGCATCTTCAACCTTGATGGATTGCT 931  
 Qy 892 GCTTATCTCTGTAGATCTTACACATCTTTTAAAGCAGGAAATTTGCTTTTGTTCAAA 951  
 Db 932 TCTTACCTCTGTCTTCTACAGGSCCTTTGAATCTAACAAGTGCTTCCCTGCCAGAT 991  
 Qy 952 GAAGGTGGCCAAATATGGTTCATTTTCTGATAGATTTCATTTCAAAAATATGAAGACT 1011  
 Db 992 CAAGGTGCCACAGATGGGTCACTATCTGATAAATTTGCTCAAGACAAGTGA---T 1048  
 Qy 1012 AATGATCACATTTATTTTAAACACAGGTCCTTTCCCATTTTCCCGTTGGAGGCAC 1071  
 Db 1049 GAGACACAGAAATACCTTCTGACACCGAGATTCACGCAATTTTGTCTCGTGGAGATAC 1108  
 Qy 1072 AAATGCTGTAACTCAGTGGAGAGCAAGTCACTCAAGAACTGTCTTCTTCTGTTGTA 1131  
 Db 1109 GGGGTTTCTATAACATTTGTCTGGGAAAAGAGCACT---GGTCAGGCTTAAAGTTGCTTTG 1165  
 Qy 1132 GCGGGGCAATTTGGGAAACCTGGGAGTTTGCATTGTGAGGAAACTTTGAGCCAGGC 1191  
 Db 1166 TTTGGAAGTAAAGGAAATACATCAATCAATTAATCTTCAAGGGATTTCTCAACACAGGC 1225  
 Qy 1192 ATGACTTACACAAATTAATCGATCAGATGTTAACTGTTGAAACATTTCAAGTGTTCAG 1251  
 Db 1226 TCTACTCAATCCAAATGAGTTTGTGCAAGCTTGTATGTTGGAACAATTTGAGAAAGTCAAG 1285  
 Qy 1252 TTTCACTGGAAGAAACATTTGTTGCAAGTCTCAGATTAAGTTGGGAGCAGAAATGTTG 1311  
 Db 1286 TTTCTTTGGAATAACAACGTTGGTAAACCCAACTTTCCCAAGTGGTGCAGCCAGATC 1345  
 Qy 1312 ATAAATACATCTGGGAATATGGATATAAATCTACTTCTGTAGCCAAAGACATTTATGGA 1371  
 Db 1346 ACCGTCAAAAGGAGAGAGGAGAAACAGTGCACAGCTTCTGCAGCGAAAGCACCGTGG 1405  
 Qy 1372 CTAATATTTCTCAGAACCTGAAACCATGCTA 1403  
 Db 1406 GAGGATGTTCTGCTCACCTCAGCCCTGTTA 1437

RESULT 8  
 AX337813  
 LOCUS 1481 bp DNA linear PAT 09-JAN-2002  
 DEFINITION Sequence 8322 from Patent WO0194629.  
 ACCESSION AX337813  
 VERSION AX337813.1 GI:18128532  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,  
 Horrigan, S., Soppet, D.R. and Weaver, Z.

LE Cancer gene determination and therapeutic screening using signature

Gene sets

Patent: WO 0194629-A 8322 13-DEC-2001;

Avalon Pharmaceuticals (US)

Location/Qualifiers

1. 1481

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

COUNT 399 a 371 c 379 g 332 t

N

ry Match 27.5%; Score 386.2; DB 6; Length 1481;

ic Local Similarity 57.1%; Pred. No. 5.3e-83;

ches 807; Conservative 0; Mismatches 588; Indels 18; Gaps 5;

1 ATGCTTGAATTTGGATTGTTGATCTTCTTTGGCAGATCAAGAGGAAAGAGTT 60

20 ATGCTGATCTTCTGGCAATCAACACTTTTCTGCTGGGAGCAGCCAAAAGGAAAGAGTT 79

61 TGCTATGAAGAGTTAGGCTGTTTCAAGATGTTTACCATGACAGGAGACTTCTCAACA 120

80 TGCTATGAGAGCTCGGGTCTTTCTGACACTGAGCCCTGGGGCGGACAGCAATCAGG 139

121 GAGTGTGAGTTTACCCCTGGTCTTCAGAGAGATAAACACTGTTTCTGCTCTACACT 180

140 CCCCTGAAATTTCTCCCTGGAGCCCTGAGAGATCGGCACCCGCTTCTGCTGTACACC 199

181 ATACACATCCCATGCTATCAGGAGATCAGTGGGTTAATTTCTCAACTATCAAGCC 240

200 AATGAACCCCAACAACTTTCAATTCCTCTCTGATCCATCAACATTCAGGCA 259

241 TCATATTTTGAACAGACAAAGATCAACCGTATCAACATAGCTGG-----ATGGAACA 294

260 TCAATTTTCAATGGACAGAAAGACCCGGTTCATCATCGATGCTTCATAGCAAGA 319

295 GATGCAATGCGAGAGACATCTGCAATGTCTGTACAGCTGGAAGATATAAATTGC 354

320 GATGAGAGCTGGGTGACAGACATGTGCAAGAACTGTTGAGGTGAGAGGTGAATGC 379

355 ATTAATTTAGATTGGATCAACGGTTTACGGG---AATACATCATGCTGTAAACATCTC 411

380 ATCTGGTGGACTGGAAGAGGGTCTCCAAAGCCACCTACACAGGCTGCCAACACGTG 439

412 CGTGTGTTGCTGAGTGGCTTATTTTATGATGTTCTCATGAAATTTGAATAT 471

440 CGAGTGTGGCGCCAGGTGGCCAGATGCTGACATCCTCTTGACAGATAGCTAC 499

472 TCCCTCTTAAAGTGCACTTGAATGGCCACAGCTTGGGACACACCTGGCTGGGAGCT 531

500 CCCCTTCCAAGTTCACTTATGGCCACAGCTGGAGCCACCGTGGAGAGGCA 559

532 GGGTCAAGGATACAGGCTTGGAGAAATACCTGGGTGAGCCAGCTGGGCCATTTTC 591

560 GGAAGCAAGACTCCAGGCTGAGCAGATTTACAGGTTGGATCTCTGAGAAGCAAGTTTC 619

592 CACACACTCCAAAGGAGTCAAGCTAGACCCCTCGATGCCCACTTTGTTGACGTTAT 651

620 GAGATCTCTGAGAGGTGCGACTTATGATCCCTCTGATGCTGACTTTGTTGATGAT 679

652 CATACAAATGAGCTCGCATCCTTTTGTGCTGGTGTGGTGGAAACCATGATGCTGTGT 711

680 CACAGGATGAGCTCCCTGATCCCTGATCCCTTCTGGGTTTGGAAACCAACAGATGGGT 739

712 CATCTTGACTTTTACCAATGAGGAGGAGCAGATGCCAGATGTGAAGACTTAATTACA 771

740 CATCTTGACTTTTCCCAATGAGGAGGAGCAGATGCCAGATGTGAAGACTTAATTACA 796

772 CTTTACTGAAATTTTAACTTCAATGCTTACAAAAGAAATGGCTTCTTTTGTACTGT 831

797 CTGTCTCAGATCGTGGATCTAGATGGCATCTGGCGGGAACCCGGGACTTTGTGGCTGC 856

832 AACCATGCCGAGTTATCAATTTTATGCTGAAAGCATCTTAACTCTGATGATTTAT 891

Db 857 ATACCTTAAGAGCTACAGTATTACTTGGAAAGCATCTCAATCCGATGGGTGTGCT 916

Qy 892 GCTTATCCTTGTAGATCCTACACATCTTTTAAAGCAGGAATTTGCTTTTGTTCCAA 951

Db 917 GCATATCCTGCACTTCTTACAGTCTTTGAGTCTGACAGTCTTCCCGTGTCCAGAT 976

Qy 952 GAAGGTTGCCCAACAATGGGTCAATTTTCTGATAGATTTTCACTTCAAAAATATGAAGACT 1011

Db 977 CAGGATGCCACAGATGGTCACTATGCTGATAAATTTGCTGGCAGCAAGTGA---A 1033

Qy 1012 AATGATCAGATATTTTAAACACAGGTCCTTTTCCCAATTTGCCGTTGGAGGCAC 1071

Db 1034 GAGCAGCAGAAATTTCTTGAACACAGGAGGCTAGCAATTTGCTGCTGGAGATAT 1093

Qy 1072 AATATGCTGTTTAAACTCAGTGGAGCGAAGTCACTCAAGGAAGTCTTTTCTTCGTGTA 1131

Db 1094 GGGGTTTCCATCAGTCTGCTGGAAGACAGCCACT---GGTCAGATCAAAAGTTGCTTTG 1150

Qy 1132 GCGGGGCAATTTGGGAAACTGGGGAGTTTGCATTTGTCAGTGGAAAACCTTGACCCAGGC 1191

Db 1151 TTTGAAATPAAGGAAACACTCACCAGTACAGCATCTTTCAGGGGATTTCTCAAAACAGGC 1210

Qy 1192 ATGACTTACACAAATTAATCGATCAGATGTTAACTGTTGAAGCAATTTACAAGTGTTCAG 1251

Db 1211 TCAACCATCTCTGAGTTGATGCAAGCTGGATGTTGGAACAATTTGAGAAGTCAAG 1270

Qy 1252 TTATCTGAAAACAACTTTGTTGAAGATTTCTGAAATAAGTTGGGAGCAGAAATGGTG 1311

Db 1271 TTTCTTTGGAATAAACAATGTGATAAATCCAACTTCCCAAGTGGTGCCCAAGATC 1330

Qy 1312 ATAAATACATCTGGGAATATGATATAAATCTACTCTCTGTAGCCAGACATTTAGGGA 1371

Db 1331 ACTGTCAAAGGAGAGAGAGACAGTGTACACTTCTGTAGCGAAGACACAGTGGG 1390

Qy 1372 CCTAATATTTCTCAGAACCTGAAACCATGCTAA 1404

Db 1391 GAAGACACGCTGCTCACCTCACGCTGCTGCTAA 1423

RESULT 9  
HUMLRP1P 1481 bp mRNA linear PRI 07-JAN-1995  
LOCUS Human pancreatic lipase related protein 1 (PLRP1) mRNA, complete  
DEFINITION cds.

ACCESSION M3283

VERSION M3283.1 GI:187229

KEYWORDS lipase related protein 1.

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1481)

AUTHORS Gillier, T., Buchwald, P., Blum-Kaelin, D. and Hunziker, W.

TITLE Two novel human pancreatic lipase related proteins, hPLRP1 and hPLRP2. Differences in colipase dependence and in lipase activity

J. Biol. Chem. 267 (23), 16509-16516 (1992)

JOURNAL MEDLINE

PUBMED 92355622

COMMENT 1379598

FEATURES

source text: Homo sapiens pancreas cdna to mRNA.

Location/Qualifiers

1. 1481

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/tissue\_type="pancreas"

1. 1481

/gene="PLRP1"

20. 1423

/gene="PLRP1"

/codon\_start=1

/product="lipase related protein 1"

/protein\_id="AAA59532.1"



```

/db_xref="GI:187230"
/translation="MLIFWNTLPLLGAAGKEVCYVEDLGCFSPTDPWGTAIRPLKI
LPSREKIGTRFELLYTNEPNFNOILLSDPSTIEASNPNFDRKTRPIHGFDKGD
SWTDMCKLFEVEBNCICUDWKGQSQTQOANRVVUQAQVOMLILLITSEY
PSKVLHLGHSUGAHVAGEAGSKTPGSLRTGLDPVEASFESTPEVRLLDPDADFV
VHTDAALPIPLPGFTNQWGHLDFFPENGESMPGCKKQALSIQVLDLGNAGTRFD
VACNHLRSYKYLYLESILNPDGFAAYPCTYSKFSFSDKCFPCPDGQCPOMGHYADKFA
RTSEQQKPLNTGASNPFAWRVGSITLSGRATGQIKVALFGNKNHGYISIFRG
ILKPGSTHSYEPDAXLDVDTIEKVKFLWNNVINPTLPKVGATKIYQKGEETVNF
CSEDTVREDTLLTLTPC"
sig_peptide 20..70
/gene="PLRPI"
mat_peptide 71..1420
/gene="PLRPI"
/product="lipase related protein 1"
polyA_signal 1458..1464
/gene="PLRPI"
polyA_site 1481
/gene="PLRPI"
COUNT 399 a 371 c 379 g 332 t
N

rv March 27.5%; Score 386.2; DB 9; Length 1481;
t Local Similarity 57.1%; Pred. No. 5.3e-83;
ches 80%; Conservative 0; Mismatches 588; Indels 18; Gaps 5;

1 ATGCTTGGAAATTGGATTGTCATCTTTGTTTGGCCATCAAGAGGAAAGAAAGTT 60
20 ATGCTGATCTTCTGGACATCAACATTTTCTGCTGGGAGCAGCCAAAGAAAGAAAGTT 79
61 TGCATGAAGGTTAGGTTTCAAGATGGTTTACCATGGACGACGACTTCTCAACA 120
80 TGCATGAGGACCTCGGTTCTTCTGACATGAGCCCTGGGCGGACAGCAATCAGG 139
121 GAGTTCGTAGTTTACCTCGTCTCAGAGAGATAACAACCTCGTTTCTGCTCTCACT 180
140 CCCTGAAATTTCTCCCTGGAGCCTGGAAGATCGGCAACCCGCTTCTGCTGTA 199
181 ATACAAATCCCAATGCCCTATCAGGAGATCAGTGGGTTAATCTTCAACTATCCAAGCC 240
200 AATGAAACCCCAACAACTTCAAAATTTCTCTCTCTCTGATCCATCAACAAATGAGGCA 259
241 TCATATTTTGGACAGACAGATCACCGGTATCAACATAGCTGG-----ATGAAACA 294
260 TCAATTTTCAATGAGACAGAACCCCGTTTCATCTCCATGCTTCATAGACAAGGA 319
295 GATGCAAAATGGCAGAGACATGTGCAATGTGTTGCTAGAGCTGGAGATATAAATGGC 354
320 GATGAGAGCTGGGTGACAGACATGTGCAAGAACTGTTTCGAGGTGGAGAGGTGAATGC 379
355 ATTAAATTTAGATTGGATCAAGCTTCAAGG---AATACATCCATGCTGAACAATCTC 411
380 ATCTGCTGACTGGAGAGAGGCTCCCAAGCCACTACACACAGCTGCCCAACAGCTG 439
412 CGTGTGTTGCTGCTGAGCTGGCTTATTTATTTATGTTTCTCATGAAATAATTTGAATAT 471
440 CGAGTGTGGCGCCAGCTGGCCAGATGCTCGACATCTCTTGACAGAGATATAGCTAC 499
472 TCCCTTCTAAATGACATTTAGTGGCCACAGCTTGGGAGACACACCTGGCTGGGAGCT 531
500 CCCCTTCTCAAAAGTTCACTCTATGAGCCACAGCTGGGAGCCACCTGGCTGGAGGCA 559
532 GGTCAAGGATACAGGCTTGGAGAAATACTGGTTGGACCCAGCTGGGCCATTTTTC 591
560 GGAAGCAAGACTCAGGCTGAGCAGATTAAGGGTTGGATCCCTGTAGAGCAAGTTTC 619
592 CAACAACATCCAAAGAGTCAGCTAGACCCCTGGATGCCAACTTTGTTGACGTTAT 651
620 GAGATGACTCCTCAAGAGGTGCGACTTGTATCCCTCTGATGCTGACTTTGTTGATGCT 679
652 CATCAAAATCGACTCGCTCTCTTTGAGCTTGTGTTGGAACCAATGATGCTTGTGCT 711
680 CACAGGATGAGCTCCCTGATCCCATTTCTTGGGTTTGGAAACCAACAGATGGGT 739

```

```

QY 712 CATCTTGACTTTTACCACAAATGAGGAGGACACATCCAGGATGTGAAGACTTAATATACA 771
DB 740 CATCTTGACTTTTCCCCAATGAGGAGGAGAGCATCGCGGATGCAAGAA--GAATGCC 796
QY 772 CTTTACTGAAATTTAATCTCAATGCTTTACAAAAGAAATGGCTTCTCTTTTGACTGT 831
DB 797 CTGCTCAGATCGTGGATCTAGATGGCATCTGGGCGCGGAAACCCGGGACTTTGTGGCTGC 856
QY 832 AACCATGCCCGAAGTTTCAATTTTATGCTGAAAGCATCTTAACTCTGATGATTTAT 891
DB 857 AATCACCTAAGAGCTACAGTATTAATCTGGAAGCATCTCAATCCGATGGTTTGTCT 916
QY 892 GCTTATCTCTTGTAGATCTACATCTTTTAAAGACGAAATGCTTCTTTTGTTCAAA 951
DB 917 GCATATCCTCTGCACTTCTTCAAGTCTCTTGAAGTCTGCAAGTCTTCCGCTGTCCAGAT 976
QY 952 GAAGCTTGGCCCAACAATGGGTCATTTGCTGATAGATTTTCATTTCAAAAATATGAAGACT 1011
DB 977 CAAGATGCCACAGATGGGTCACTATGCTGATAAATTTGCTGCGAGGACAAGTGA--A 1033
QY 1012 AATGATCACATTAATTTTAAACACAGGGTCCCTTTCCCATTTGCCCTTGGAGGCAC 1071
DB 1034 GAGCAGCAGAAATTTCTTCTTGAACACAGGAGAGCTAGCAATTTCTGCTGGAGATAT 1093
QY 1072 AAATGTCTGTTAACTCAGTGGAGGAGTCACTCAAGAACTGCTTTTCTTCTGCTA 1131
DB 1094 GGGGTTTCCATCACACTGTCTGGAAGAACAGCCACT--GGTCAGATCAAAGTTGCTTTG 1150
QY 1132 GCGCGGGCAATTTGGGAAACTGGGGAGTTTGCCTATGCTAGTGGAAAACTTGAGCCAGGC 1191
DB 1151 TTTGAAATAGGAAACACTCACCAGTACAGCATCTTCAGGGGATTTCTCAAACAGGC 1210
QY 1192 ATGACTTACAAAAATTAATCGATGAGATGTTAACTGTTGGAACATTAAGTGTTCAG 1251
DB 1211 TCAACCCATTTCTATGATTTGATGCAAGCTGATGTTGGAAACAATTTGAGAAAGTCAAG 1270
QY 1252 TTTCACTCGAAAAACATTTCTTTGAAGATTTCTCAGAAATAGTTGGGAGCAGAAATGGTG 1311
DB 1271 TTTCTTTTGGATTAACATGTGATTAATTAACACCTTCCCAAGTGGTGGCCACCAATC 1330
QY 1312 ATAAATACATCTGGAAATATGGATATAAATCTTCTGTAGTCCAGCAATTTATGGGA 1371
DB 1331 ACTGTGCAAAAGGAGAGAGACAGCAGTGTACAACTTCTGTAGCGAAGACACAGTGGG 1390
QY 1372 CCTAATATTTCCAGAACCTGAAACCATGTAA 1404
DB 1391 GAAGACACGCTGCTCACCTCACGCTTGTCTAA 1423

RESULT 10
BC025784 1513 bp mRNA linear PRI 11-MAR-2002
LOCUS IMAGE:5226231, mRNA, complete cds.
DEFINITION BC025784
ACCESSION BC025784.1 GI:19343957
VERSION BC025784.1
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1513)
AUTHORS Strausberg, R.
DIRECT SUBMISSION National Institutes of Health, Mammalian
TITLE Submitted (06-MAR-2002)
JOURNAL Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.

```



CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINLU)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland.  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc.mcgenhgr1.nih.gov](mailto:nisc.mcgenhgr1.nih.gov)  
Shevchenko, Y., Wetherby, K.D., Sackett-Strom, S.M.,  
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
Dieciuch, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Snyder, B., Stantirip, S., Thomas, P.J.,  
Tingston, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,  
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: TRAK Plate: 49 Row: n Column: 13  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5453919.

URES  
source

```
1. localid="guar1000005"
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocustID:5407"
/db_xref="taxon:9606"
/clone="WGC:3434 IMAGE:522631"
/tissue_type="Pancreas, Spleen,
clone lib="NIH MGC_120"
/lab_host="PH108"
/notes="Vector: pCMV-SPORT6"
```

CDS

```

/notes vector: pcmy3-SPOR16
28. .1431
/codon_start=1
/product="pancreatic lipase-related protein 1"
/protein_id="AAH25784.1"
/db_xref="GI:19343958"
/translation="MLIFMTITFLGAAKGVKCYEDLGCSTDEPWGQTARLPLKI
LPSWCKITRELPYLNENNFQLISDSTISANFVQMGAKTRFIILHGFIDKGEY
SWYDMCKILFVEEYVNCIVCMVKQSGATQAAANYVQVGAQWMLDILTEYSY
PPSKVHLIHSILCAHVAGSAGSTPGLSRIITGLDPEASFESEETVELVLPDSADPVD
VHTIDAAPILPLFGTGTQOQMGHLDFPFGNGSPMGCKKNALSOILVDLGTWAGTRFD
VACNHLRSVKVYLETSLNEDGAVAPVCTSYKSFESDKCFPCPDQGCQMGHYAKPAG
RTSEEQKSFPLNTGEASNPFRWYGVSTILSGRTATGQIKVALFGNKGNTQKSYIFNG
ILKPGSTYFSEDAKLDDITDKVFLMNNVINFLPKVGAITIKVQKEGKVTYVNF
CSBETVREDTTLTLTPC"
372 c 382 t
372 c 332 t

```

COUNT  
IN

```

ery Match      27.5%; Score 386.2; DB 9; Length 1513;
st Local Similarity 57.1%; Pred. No. 5.3e-83;
tches 807; Conservative 0; Mismatches 588; Indels 18; Gaps 5;

1  ATGCTTGGAAATTTGGATTGTTGCAATCTCTTGCTTTGGCACATCAAGAGGAAAGAAATTT 60
28  ATGCTGATCTTCTGGACAATCAACATCTTTCTCTGCTGGGAGCAGCCAAAGGAAAAGAAATTT 87
61  TGCTATGAAAGGTTAGGGTGTCTTCAAAGATGGTTTACCATTGGACACGAGACTTTCTTCAACA 120
88  TGCTATGAGGACTTCGGTGCTTTTCTGACACTGAGCCCTGGGGGGGAGACGAACTCAGG 147
121  GAGTTGGTAGTGTTTACCCCTGGTCTCCAGAGAAGATAAACACTCGTTCTCTGCTCTACACT 180
148  CCCCAGAAATTTCTCCCTGGAGCCCTGAGAAGATCGGCACCCGGCTCTCTGCTGTACACC 207
181  ATACACAAATCCCAATGCTTATCAGGAGATCACTGGGGTTTAATCTTCTCAACTATCCAGGCC 240
208  AATGAAACCCCAACAACTTTCAAATTCCTCCTCTCTGATCCATCAACAAATTTGAGGCA 267
241  TCATATTTTGGACACAGACAGATCAACCCGTFATCAACATAGCTGG-----ATGGAARACA 294
268  TCAAAATTTTCAAATGGACAGAAAGACCCCGTTTCATCATCCAATGGGTTTCATAGACAAGGA 327
295  GATGGCAAATGGCAGAGAGACATGTGCAATGTGTGCTACAGCTGGGAGAGATTAATATGC 354

```

338 GATGAGAGCTGGGTGGACAGACATGTGCAAGAAACATGTTTCGAGGTGGAGGGTGAATGCG 387

355 ATTAATTTAGATTGGATCAACGGTTTCAACGGG---AATACATCCATGCTGTGTAAACAATCTC 411

368 ATCTGCGTGGACTGGAGAAGGCTCCCAAGCCACTACACAGGCTGCCAACACAGTG 447

412 CGTGTGTTGGTGTGAGGTGGCTTATTTTATGTGTTCTCATGAAAAAATTTGAATAT 471

448 CGAGTGTGGGCGCCCGAGGTGGCCAGATGCTCGACATCTCTTTGACAGAGTATAGCTAC 507

472 TCCGCTCTAAAGTGACATTTGATTGGCCACAGCTTGGGACACACCTGCGCTGGGGAAGCT 531

508 CCCCTTCCAAAGTTTCACTTCATTGGCCACAGCTTGGAGCCCGCTGGGCTGGAGAGGCA 567

532 GGGTCAAGGATAACAAGCCCTTGAAGATAAATCTGGGTTGGACCCAGCTGGGGCCATTTTTC 591

568 GGAAGCAGAGACTCCAGGCTTGAGCAGGATTAACGGGTTGGATCTCTGTAGAAGCAAGTTTC 627

592 CACAACACTCCAAAGAAAGTCAGGCTTAGACCCCTCGGATGCCAATTTGTTGGAGCTTATT 651

628 GAGAGTACTCTGAGAGAGTGCGACTTGATCCCTCTGATGCTGACTTTGTTGATGTGAT 687

652 CATACAAATCGAGCTCGCATCTCTTTGAGCTTGGTTGGAACCAATGATGCTTGTGGT 711

688 CACAGGATCGAGCTCCCTGATCCATCTTTGGGTTTTTGGAAAGCAACACAGATGGT 747

712 CATCTTGACATTTACCAAAATGGAGGAGACATGCCAGAGTGTGAAGACTTAAATTACA 771

748 CATCTTGACTTCTTCCCCAATGAGGAGAGACATGCCGGGATGCAAGAA---GAATGCC 804

772 CTTTATCTGAATTTAACTTCAATGCTTACAAAAAGAAATGGGTTCTCTTTTGACTGT 831

805 CTGTCTCAGATCTGTGGATCTAGTGCATCTGGCGGGAACCCGGGACTTTGTGGCTTG 864

832 AACCATGCCGAGATTATCAATTTATGCTGAAACATCTTAACTCTGATGCAATTTATT 891

865 AATCACCTAAGAGACTCAAGTATTACTTTGAAAGCATCTCAATCCCGATGGGTTTGTCT 924

892 GCTTATCTGTAGTCTTACATCTTTTAAAGCAGAAATTGCTCTCTTTTGTTCCTAAA 951

925 GCATATCTCTGCATCTCTTACAAGTCTTTGAGTCTGACAGTGTCTCCGCTGCCAGAT 984

952 GAAGGTTGCCCAACAATGGGTCAATTTGCTGTATAGATTTCACTTCAAAAATATGAAGACT 1011

985 CAAGGATGCCCAAGATGGGTCACTATGCTGATAAATTTGTGGCAGGACAAAGTGA---A 1041

1012 AATGGATCATATTATTTTAAACACAGGCTCCCTTTCCCATTTGCCGTTGGAGGCAC 1071

1042 GAGCAGCAGAAATCTCTTTTGAACACAGGAGGCTAGCAATTTGCTCGCTGGAGATAT 1101

1072 AAATTTGTCTGTAACTCAGTGGAGCGAAGTCACTCAAGAACTGTCTTTCTTCTGTGTA 1131

1102 GGGGTTTCCATCACTGTCTGGAAGACAGGCACT---GGTCAGATCAAAAGTTGCTTTG 1158

1132 GGCGGGCAATTGGGAAACTGGGAGTTTGGCATTTGCTAGTGGNAACTTTGAGCCAGGC 1191

1159 TTTGGAATTAAGGAAACACTCACAGTACAGCATCTTCAAGGGGATTTCTCAAAACAGGC 1218

1192 ATGACTTACAAAAATTAATTCGATGCAGATGTTACGTTGGAACAATTACAAGTGTTCAG 1251

1219 TCACCCATTTCTATGATTTGATGCAAGCTGGATGTTGGACAAATTTGACAAAGTCAAG 1278

1252 TTCACTTGGAAAAACAATTTGTTTGAAGATTTCTCAGANATAGTTGGGAGCAGAAATGGTG 1311

1279 TTTCTTTGGAAATACATGTGTGATAAATCAACCCCTCCCAAGTGGGTGCCACCAAGATC 1338

1312 ATAAATACATCTGGAAAAATGGATATAAATCTACCTTCTGTAGCCAGACATTAATGGGA 1371

1339 ACTGTGCAAAAGGAGAAAGGAAGACAGTGTACAACTTCTGTAGCGAAGACACACAGTGGG 1398

1372 CCTAATATTTCCAGAACCTGAAACCATGCTAA 1404

1399 GAAGACACGCTGCTCACCCCTCAGCCCTGCTAA 1431

T 11

ECPL 1443 bp mRNA linear MAM 22-JUL-1993  
 ION E.caballus mRNA for pancreatic lipase.  
 SION X66218  
 ON X66218.1 GI:1063  
 RDS lipolytic enzyme; pancreatic lipase.  
 E Equis caballus (horse)  
 ANIM Equis caballus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 1  
 ENCE Kerfelec, B., Foglizzo, E., Bonicel, J., Bougis, P. E. and Chapus, C.  
 HORS Sequence of horse pancreatic lipase as determined by protein and  
 LE CDNA sequencing. Implications for p-nitrophenyl acetate hydrolysis  
 by pancreatic lipases  
 Eur. J. Biochem. 206 (1), 279-287 (1992)  
 RNAL  
 LINE 92267022  
 BMED 1587279  
 ENCE 2 (bases 1 to 1443)  
 HORS Kerfelec, B.  
 LE Direct Submission  
 RNAL Submitted (22-JUL-1992) B. Kerfelec, CNRS, CBM, 31 Chemin Joseph  
 RES Aguer, 13402 Marseille Cedex 9, FRANCE  
 Location/Qualifiers  
 1. .1443  
 /organism="Equis caballus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9796"  
 /tissue\_type="pancreas"  
 /clone\_lib="pancreatic cDNA"  
 1. .1443  
 /gene="lipase"  
 <1. .1386  
 /gene="lipase"  
 /codon\_start=1  
 /product="pancreatic lipase"  
 /protein\_id="CAA46961.1"  
 /db\_xref="GI:1064"  
 /db\_xref="SWISS-PROT:P29183"  
 /translation="MTLSLLGAVGVNVCYVERLGFSDSPWAGIVERPLKILPWSP  
 EKVNTPFLYTHNENPDQEIIVADPTLQSNFNTGKTRFIIHGFDKGEESWLSM  
 CQMFVESVNCICVDKSGSTAIQSQSNRIVGAEVAYLVGLSSDYSFSPNVH  
 IIGHLSHGAAGEAGRRINGAVRITGLDPAEPQPGPELVRLDPSDAQFVDVHTD  
 IAPFIPNLGFGMSQTAGHLDFFPNKGKMPQCKNLSQIVDIDGIWQGRTRDPAACNH  
 LRSKYTDSIILNPQDFAGFCASYSDFSTANKFCPCSEGPCQMGHVAIDRFPQRTKV  
 GOLFYLTGDSNFAWRVRVDVTLGKVKVTHLVSLFGNKGNSROVEIFQGLKPD  
 NTVSNFSDVSDVGLDEKVKPFWYNNVINLTLPKVGASKITVERNDGSGVNFCSSETV  
 REDVLLTLTAC"  
 <1. .36  
 /gene="lipase"  
 37. .1383  
 /gene="lipase"  
 /product="pancreatic lipase"  
 /evidence=experimental  
 37. .1044  
 /gene="lipase"  
 /note="N-terminal domain"  
 /evidence=experimental  
 748. .822  
 /gene="lipase"  
 /note="flap"  
 1045. .1356  
 /gene="lipase"  
 /note="C-terminal domain"  
 /evidence=experimental  
 1422. .1428  
 /gene="lipase"  
 394 a 324 c 356 g 369 t  
 COUNT  
 N

Query Match 26.8%; Score 376.2; DB 4; Length 1443;  
 Best Local Similarity 58.5%; Pred. No. 1.4e-80;  
 Matches 812; Conservative 0; Mismatches 548; Indels 27; Gaps 8;  
 QY 29 TGTCTTTGGCACATCAAGAGGAAAGAGTTTCTATGAAGTTAGGTAGGTTCATCAAG 88  
 DB 14 TGTCTCTGGGAGCAGTAGTAGGAAACGAAAGTTTCTATGAGAGACTCGGCTGCTCAGTG 73  
 QY 89 ATGGTTTACCATTGACACAGGAGACTTCTCAACAGAGTTGGTAGTTTACCCCTGGTCTCCAG 148  
 DB 74 AGCATTCGCGTGGCGGGGATTGTGGAAGACCTCTCAAAATATTGCTCTGCTCCAG 133  
 QY 149 AGAAGATAAACAACATGTTTCTGTCTACATAACATACATCCCAATGCTCTATCAGAGA 208  
 DB 134 AAAAGGTCAACACCGCTTCTCTGTACACTAACAGAGAACCACCACTTTCAAGAAA 193  
 QY 209 TCAGTGGCGTTAATCTTCAACTATCAAGCTCATATTTTGGAAACAGACAGATCACCC 268  
 DB 194 TCGTTGCGAG---ATCCTTCAACTATCCAGACTCAATTTCAACAGAGTAGAAGAAACCC 250  
 QY 269 GTATCAACATAGCTGGAT-----GGAAAACAGATGCCAATGGCAGAGACATGTGCA 322  
 DB 251 GCTTTATTATTATGATTATCATAGCAAGGGAGAGAAAGCTGGCTGTCCACTATGTGCC 310  
 QY 323 ATGTTTCTGTACAGCTGGAAGATATAATTGCATTATTTAGATTGGATCAACGGTTTAC 382  
 DB 311 AGAACATGTTTAAGGTGGAAGAGTGGAAGTGTGAACTGTA TCTGTGTGAGCTGGAAAAGTGGCTCCC 370  
 QY 383 G---GGAATACATCATCTCTGTAAACAATCTCCCTGTTTGTGTGTGCTGAGGTGGCTTATT 439  
 DB 371 GCACGCGCATCTCAGAGCTTCAAGAACGTCCTGGATTGTGGGGGCGAAGATGCGCATATC 430  
 QY 440 TTATTGATGTTCTCATGAAAAAATTTGAATATTCCTTCTTAAAGTGCATTTGATTGGCC 499  
 DB 431 TTGTTGGAGTTCTTCAGTCATCATTTGCACTACTCTCTCTCAACGTCACATCATTTGGCC 490  
 QY 500 ACAGCTTGGGAGCACACCTGCTGGGGAAGCTGGTCAAGGATACCAAG---CCTTGGAA 556  
 DB 491 ATAGCTGTGTTCTCACGCTGTGGGAGGCTGGAAGGAGACCAATGGGGCGCTGGAC 550  
 QY 557 GAATAACTGGTGTGACCCAGCTGGGCCATTTTTCACAACACTCCAAAGAGATCAGGC 616  
 DB 551 GGATCACAGGGTTGGATCCAGCTGAACCTTCTTTCAGGGGCACACCCGAAATAGTCCGAT 610  
 QY 617 TAGACCCCTCGATGCCCACTTGTGTGAGTTATTACAAATGCAAGCTCGCATCTCT 676  
 DB 611 TGGACCCCGAGGATGCCCAATTTGTGGATGTAATTCACAGACATTTGCCCTTCAACC 670  
 QY 677 TTGAGCTTGGTGTGGAAACCATTTGATGTTGTGTGCTCATCTTGACTTTTACCCAAATGGAG 736  
 DB 671 CCAACTTGGGGTTTGGAAATGAGCCAAACTGCAGGCCACCTGGATTCTTTCMAAATGGAG 730  
 QY 737 GGAAGCAGATGCCAGGATGTGAAGCTTAATTACCTTTACTGAATTTAACTTCAATG 796  
 DB 731 GAAAAGAAATGCCGGGATGTCAAAAGAACGTT---CTCTCTCAGATTGTTGACATAGACG 787  
 QY 797 CTTACAAAAAGAAATGGCTTCTTCTTTGACTGTAAACCATGCCCGGAAGTTATCAATTTT 856  
 DB 788 GGATCTGGCAGGAACCTCGTACTTTGGCGCTGTATCACTTAAGAAGCTACAAGTATT 847  
 QY 857 ATGCTGAAGACATTTTAAATCTGTAGCATTTATTTGCTTATCTTGTAGATCTTACACAT 916  
 DB 848 ACATGATAGCATCTCTCAACCCCTGATGGCTTTGCTGGATTCTCTTGTGCTCTTTACAGTG 907  
 QY 917 CTTTAAAGCAGGAAATTTGCTTTTCTTCCAAAGAGGTTTGGCCCAACATGGGTCAAT 976  
 DB 908 ATTTCACTGCAAAACAGTGCTTCCCTGTTTCAAGTGAAGGCTGCCACAGATGGGTCAAT 967  
 QY 977 TTGCTGATAGATTTCACCTTCAAAAATATCAAGACTAATGGAATCACTATTTTATAACA 1036  
 DB 968 ATGCTGATAGATTTCCTGGAG---ACAAAAGAGTAGGGCCAGCTATTTTATCTAAACA 1024

1037 CAGGTCCTTTCCCATTTGCCGTTGGAGCACAATAATTGCTTTAACTCAGTGAA 1096  
 1025 CCGTGATGCCACCAATTTTCCCGTTGGAGGTATAGGTAGATCTCAGCTCTGGAA 1084  
 1097 GCGAAGTCACTCAAGAACTGCTTTCTTCGTTGGTGGCGGGCAATTCGGAAACTGGGG 1156  
 1085 AGAAGGTTAC---AGGACACGCTGCTAGTTCTTCTTTGTTGGAAATAAGGGAACCTAGGC 1141  
 1157 AGTTTGCATTGTCAGTGGAAACTTGGCCAGGATGACTTTACACAAAATAAATCGATG 1216  
 1142 AGTATGAATTTTCCAGGGCACTCTCAACACAGACATATTTATTTCCAAATGAATGACT 1201  
 1217 CAGATGTTAACTGGAAACATPACAAAGTGTTCAGTTCACTCTGGAAGAAACATTTGTTG 1276  
 1202 CAGATGGAAGTTCAGATTTGGAGAGGTTAAATTTATTTGGTACAAACATGTGATCA 1261  
 1277 AGATTTCTAGATAGTTGGGAGCAGAAATGGTATATAATACATCTCGGAATATGAT 1336  
 1262 ACCTAATCTACCCCAAGTAGGAGCATCCAGATCAGTGAA---AGAACGATGAA 1318  
 1337 ATAAATCTACTTCTGTAGCCAGACATTTATGGACCTTAATTTCTCCAGAACCTGAAAC 1396  
 1319 GTGTGTTCAACTCTGTAGTGAAGAACCGTGAGGGAAGATGTTCTGCTCACTCTCACTG 1378  
 1397 CATGCTA 1403  
 1379 CGTGTTA 1385

J. 12

274  
 HITON Mus musculus pancreatic lipase related protein 1 mRNA, complete

AF061274 1620 bp mRNA linear ROD 10-MAY-1999

AF061274

AF061274.1 GI:3108174

Mus musculus (house mouse)

Mus musculus

Phylum: Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1620)

Remington, S.G., Lima, P.H. and Nelson, J.D.

Pancreatic lipase-related protein 1 mRNA in female mouse lacrimal

gland

Invest. Ophthalmol. Vis. Sci. 40 (6), 1081-1090 (1999)

99249620

JMED

2 (bases 1 to 1620)

Remington, S.G., Lima, P.H. and Nelson, J.D.

Direct Submission

Submitted (24-APR-1998) Ophthalmology, Regions Hospital, 640

Jackson Street, St. Paul, MN 55101, USA

Location/Qualifiers

1. 1620

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="Swiss Webster"

/db\_xref="taxon:10090"

/sex="female"

/tissue\_type="exorbital lacrimal gland"

158. 1579

/note="PLRP1"

/codon\_start=1

/product="pancreatic lipase related protein 1"

/protein\_id="AAC15774.1"

/db\_xref="GI:3108174"

/translation="MLIWIPIFLGAAAGKVCVDNLCGSDPAPWAGTAIRPLKL  
 LPWSPKINTFLYLTNPFATLQLSDFTIEASNFQARKTFPIHGFIDRGE  
 NWVDMCKMFEVENVNCICVDKRGSGTITQANNRVVGVQVAMIDILVRFNY  
 SASKVHLIGHLAGHVAAGASRTPGIITGLDPVEANFEGTPEVRILDPDADFVD  
 VIHTDAFLIPLFGTGNOMVGHDFPFNGGVMPGCKKALSOIVIDIGIWSGTRDF  
 VACNELRSKYLYESSILNPDGFAAYPCASVYDFESNKCPCPCPQGGQMGHVDKFN

CDS

NTSVEPKFFLNTGEAKNFARWRYRVSLETSGRTVQGVKVSFLGNSGNTRQCDIFRG  
 IIKGATHNSFEFDALQVGTIEKVKFLNNHVNVPSPKVGAAKITVQGBERTEHNF  
 CSEETVEDILLTLPLCKTSDTM"

BASE COUNT 422 a 415 c 410 g 373 t

ORIGIN

Query Match 26.6%; Score 373.4; DB 10; Length 1620;  
 Best Local Similarity 56.5%; Pred. No. 6.8e-80;  
 Matches 799; Conservative 0; Mismatches 596; Indels 18; Gaps 5;

QY 1 ATGCTTGGAATTTGGATTTGTCATTCTGTCCTTTGSCACATCAAGAGGAAAGAAAGTT 60  
 Db 158 ATGCTTGATTTCTTGAGCAATCCCTCTTTTCTACTGGAGCAGCTCAGGGGAAAGAGTT 217  
 QY 61 TCTATGAAAGGTTAGGTTGTTTCAAGATGGTTTACATGACACAGGACTTTCTCAACA 120  
 Db 218 TCTATGACAACTTTGGATGCTTTTCTGATGCTGAGCCCTGGGCGAGGACAGCTATCAGG 277  
 QY 121 GAGTTGGTAGTTTACCCTGCTCCGAGAGGATAAACACTCGTTTCTGCTCTACACT 180  
 Db 278 CCCCTAAACATTTCTCCCTTGGAGCCCTGAGAAGATCAACACTCGCTTCTGCTCTACCC 337  
 QY 181 ATACACAATCCCAATGCCATATCAGGAGATCAGTGGGTTAATTTCTCAACTATCCAGCC 240  
 Db 338 AATGAGAACCCAACTGCTTCCAGACTCTTCAGCTTTTCTGACCCATCGACCATGAGGCC 397  
 QY 241 TCATATTTGGAACAGACAGATCACCGTATCAACATAGCTGG-----ATGGAAACA 294  
 Db 398 TCCAATTTTCAAGTTGCCAGGAAGACTCGGTTTCATCCTCATGGCTTCATAGACAAAGGA 457  
 QY 295 GATGGCAATGCGACAGAGACATGTGCAATGTGTTGCTTACAGCTGGAAGATATAAATTGC 354  
 Db 458 GAGGAGACTGGTGGTTGACATGTGCAAGAACATGTTCCAAGTGGAGGAGGTGCACTGC 517  
 QY 355 ATTAATTTAGATTGGATCAACGGTTTCAAG---GAATACATCCATGCTGTAAACAATCTC 411  
 Db 518 ATCTGGTGGACTGGAAGAGAGGCTCTCAGACTACTTACACGCGAGGCTGCCAACAAATGTG 577  
 QY 412 CGTGTGTTGGTGGTGGTGGTCTTATTTTATGATGTTCTCATGAAAAAATTTGAATAT 471  
 Db 578 CGATGATGGTGCCAGAGTGGCTCAGATGATGACATCTTGTGAGAACTTCAACTAC 637  
 QY 472 TCCCTTTAAAGTGCATTTGATTTGGCCACAGCTTTGGAGCACACCTGGCTGGGGAAGCT 531  
 Db 638 TCTGCTTCCAAAGTCCACCTCATTTGGCCACAGCTTAGAGGCCACGTCGCGAGGGAAGCA 697  
 QY 532 GGCTCAAGGATACCGAGGCTTGGAAAGATAACTGGGTTGGACCCAGCTGGGCCCATTTTTC 591  
 Db 698 GGAAGTCGAGCTCCAGGCTTTGGCAGGATTAAGGACTGGATCTCTGTAGAGCAAAATTT 757  
 QY 592 CACAACACTCCAAAGGAAGTCAAGCTAGACCCCTCGGATGCCAACTTTGTTGAGCTTAT 651  
 Db 758 GAGGGAACCCCTGAAGAGGTCCGGCTCGACCCCTCGGATGCTGACTTTGTTGATGTGATT 817  
 QY 652 CATACAAATGCAGCTCGCATCTCTTTGAGCTTGGTGTGGAAACCATTTGATGCTTGTGCT 711  
 Db 818 CACACAGATGCAGCTCCCTTGATCCCATTTCTTGGGCTTTGGAAACAAACCAATGGTGGG 877  
 QY 712 CATCTTGACTTTTACCCCAATGGAGGAGCAGCATGCCAGGATGTGAAGACTTAATTACA 771  
 Db 878 CACTTTGACTTTCTTCCCAACGAGGAGCAGTACATGCCCGGCTGCAAGAA---GAACGCT 934  
 QY 772 CTTTACTGAAATTTAACTTCAATGCTTACAAAAAGAAATGGCTTCTCTTTCAGCTGT 831  
 Db 935 CTGTCAAGATTTGTGACATCGATGGCATCTCGTCAGGAACCCGGGACTTTTGTGGCTTGT 994  
 QY 832 AACCATGCCCGAAGTTATCAATTTTATGCTGAAAGCATTTCTTAATCCTGATCATTTAT 891  
 Db 995 AACCACTGGAAGCTACAAGTACTACTTAGAGACATCTTAAACAGATGGTGTGCT 1054  
 QY 892 GCTTATCTTTGATGCTTACATCATCTTTTAAAGCAGGAATGCTTCTTTGTTGTTCCAA 951  
 Db 1055 GCATACCCCTGCGCTTCTCTACAGGAGCTTTGATGCTTAACAAATGCTTCCCTGCCCAGAT 1114

```
952 GAAGTTCGCCCAACAATGGGGTCATTTCCTGCTAGATATTTTCACTTCAAAAATATGAAGACT 1011
1115 CAAGCTGCCCAACAAGATGGGTCATATCTGTAAGTTT---GCCAACATACAAGTGTG 1171
1012 AATGATCACATATTTTAAACACAGGGTCCCTTTCCCATTTGCCCGTTGGAGGCAC 1071
1172 GAGCCACAGAAGTTCCTCCTGAACACAGGAGAACCAAGAACTTTGACGGCTGGAGTAC 1231
1072 AAATGCTCTGTTAACTCAGTGGAGCGAGTCACTCAAGGAATGCTTTCTTCTCGTA 1131
1232 CGTGTTCCTTGACATTCCTTGGAGGACGCTCAC---AGCCAACTCAAGTGTCTCTG 1288
1132 GCGGGGCAATTTGGGAAACATGGGGAGTTTGGCCATTTGTCAGTGGAAACATTTGAGCCAGC 1191
1289 TTTGGAATTAACGCACTACTCGCCAGTGTGATCTTCAAGGGAATTAACAAGCCCGT 1348
1192 ATGACTTACACAAATTAATCGATCGAGATGTTAACTTGGGAACATTAACAGTGTCTAG 1251
1349 GCTACGCAATCCATGAGTTTGTATGCCAAGCTTGTATGGGAACAATTTGAGAAGTCAAG 1408
1252 TTCACTCGGAAACAACTTTGTTTGAAGATTTCTCAGAAATAGTTGGGAGCAGAAATGGTG 1311
1409 TTTCTTTGGATTAACCACTGGTAAACCCAGTTTCCCAAGTGGCGCAGCCAAGATC 1468
1312 ATAAATACATCTGGGAAATATGGATATAATCTACCTTCTGTAGCCCAAGACATATGGGA 1371
1469 ACTGTGCAAAAGGGGAGGAGCGGACAGACAGACAACTTCTGTATGGAAGACCGGTGAGA 1528
1372 CCTAATATTTCCAGAACCTTGAACACCATGCTAA 1404
1529 GAAGACATCTGCTCCTCTCTTGGCTTGTA 1561

f 13
RNA
ITION RNLIPRNA 1534 bp mRNA linear ROD 19-FEB-1992
SICN R.norvegicus mRNA for lipase.
DN X61925.1 GI:56599
RDS lipase gene; triacylglycerol acyl hydrolase; triacylglycerol
lipase.
3 Rattus norvegicus (Norway rat)
ANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
BNC 1 (bases 1 to 1534)
HORS Wicker-Planquart,C. and Puigserver,A.
LE Primary structure of rat pancreatic lipase mRNA
RNAL FEBS Lett. 296 (1), 61-66 (1992)
LINE 92111786
BNED 1730292
ENCE 2 (bases 1 to 1534)
HORS Wicker-Planquart,C.
LE Direct Submission
RNAL Submitted (03-JUL-1991) C. Wicker-Planquart, CBM-CNRS, 31, Chemin
Joseph-Aiguier B.P.71, F-13402 Marseille Cedex 09, FRANCE
RES location/Qualifiers
source 1. .1534
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Wistar"
/db_xref="taxon:10116"
/clone="L9"
/tissue_type="pancreas"
/clone_lib="puc 18 cdna library"
/dev_stage="adult"
73. .1494
/EC_number="3.1.1.3"
/notes="triacylglycerol acyl hydrolase"
/codon_start=1
/product="triacylglycerol lipase"

CDS
952 GAAGTTCGCCCAACAATGGGGTCATTTCCTGCTAGATATTTTCACTTCAAAAATATGAAGACT 1011
1115 CAAGCTGCCCAACAAGATGGGTCATATCTGTAAGTTT---GCCAACATACAAGTGTG 1171
1012 AATGATCACATATTTTAAACACAGGGTCCCTTTCCCATTTGCCCGTTGGAGGCAC 1071
1172 GAGCCACAGAAGTTCCTCCTGAACACAGGAGAACCAAGAACTTTGACGGCTGGAGTAC 1231
1072 AAATGCTCTGTTAACTCAGTGGAGCGAGTCACTCAAGGAATGCTTTCTTCTCGTA 1131
1232 CGTGTTCCTTGACATTCCTTGGAGGACGCTCAC---AGCCAACTCAAGTGTCTCTG 1288
1132 GCGGGGCAATTTGGGAAACATGGGGAGTTTGGCCATTTGTCAGTGGAAACATTTGAGCCAGC 1191
1289 TTTGGAATTAACGCACTACTCGCCAGTGTGATCTTCAAGGGAATTAACAAGCCCGT 1348
1192 ATGACTTACACAAATTAATCGATCGAGATGTTAACTTGGGAACATTAACAGTGTCTAG 1251
1349 GCTACGCAATCCATGAGTTTGTATGCCAAGCTTGTATGGGAACAATTTGAGAAGTCAAG 1408
1252 TTCACTCGGAAACAACTTTGTTTGAAGATTTCTCAGAAATAGTTGGGAGCAGAAATGGTG 1311
1409 TTTCTTTGGATTAACCACTGGTAAACCCAGTTTCCCAAGTGGCGCAGCCAAGATC 1468
1312 ATAAATACATCTGGGAAATATGGATATAATCTACCTTCTGTAGCCCAAGACATATGGGA 1371
1469 ACTGTGCAAAAGGGGAGGAGCGGACAGACAGACAACTTCTGTATGGAAGACCGGTGAGA 1528
1372 CCTAATATTTCCAGAACCTTGAACACCATGCTAA 1404
1529 GAAGACATCTGCTCCTCTCTTGGCTTGTA 1561

/protein_id="CAA43927.1"
/db_xref="GI:56600"
/db_xref="SWISS-PROT:P54316"
/translation="MLTWTSLFLLGAAQGEVVDNLGCFSDAEFPWAGTAIRPLKL
LPSPKINTRPLLYTNENPTAQTQLSDPLTIGASNFQVARKTRFIHGFIDKGE
NWVDKCMKMFQVEEVNCTCDWKKGSQTYTQAANNRVVGAOVAOMDILLVKNSY
SPKVLHLIGHSGAHVAGAGSKTPELGRITIGTDFVEANFEGTPEEVRLLDPSDAFVD
VIHTDAAPLIPFLFGTQNSGHLDFPFGGSMFGCKKNLSIVDIDIGIWSGDFND
VACNHLRSYKYLESLINPDGFAAPYCSYKDFESNKFPCPDQCGQPMGHVADKFA
KSGDBPKFLPNTGEAKNFARVYRVSLLISGRMTGVQKVALFGSKGNTRYDIFRG
IIPGATHSESPDKLDVGTIEKVKFLMNNQVNPSPKVGAAKITVQGEERTYNF
CSEETVREPTLLLPFCETSDIV"
sig_peptide 73. .123
mat_peptide 124. .1491
/polyA_signal 1510. .1515
/polyA_site 1531
BASE COUNT 410 a 392 c 385 g 347 t
ORIGIN
Query Match 26.4%; Score 370.2; DB 10; Length 1534;
Best Local Similarity 56.4%; Pred. No. 4.1e-79;
Matches 797; Conservative 0; Mismatches 598; Indels 18; Gaps 5;
QY 1 ATGCTTGGAAATTTGGATTTGTCATTTCTTTCTTGGCACATCAAGAGGAAAAAGAGTT 60
Db 73 ATGCTGACCTCTGGACAGTCTCTCTTTCTTCTACTGGAGCAGCTCAGGAAAAAGAGTT 132
QY 61 TGCATATAAAGGTTAGGTTTCAAAAGATGTTTCAATGACAGACAGGACTTTCTCAACA 120
Db 133 TGCATATAAAGTGGATGCTTTCTGACGCTGAGCCCTGGGACAGGACAGCTATCAGG 192
QY 121 GAGTTGGTGTGTTTACCTGCTCCAGAGAGATCAACACTCGTTTCTGCTCTACACT 180
Db 193 CCCCTCAAACTTCCCAATGAGGCCCTGAGAGATCAACACTGCTTCTGCTGTACACC 252
QY 181 ATACACAATCCCAATGCTATCAGGAGATCAGTGCAGTTAATTTCTTCAACTATCCAAGCC 240
Db 253 AATGAGAACCCCACTGCTTTTCACTCTCAGCTTCTGACCCCATTTGACCATTTGGGGCC 312
QY 241 TCATATTTTGGACAGACAGATCACCCTGATCAACATAGCTGG-----ATGGAACA 294
Db 313 TCATATTTTCAAGTTGCCAGAGACTCGGTTTATCATCATCGCTTCTGCTGTACACC 372
QY 295 GATGCCAAATCGCAGAGAGACATGTCATGTGTGTGTCACAGCTGGAAGATATAAATTC 354
Db 373 GAGGAAACCTGGTGGTTGACATGTCAGAAACATGTTCCAAGTGGAGAGGTGAATGC 432
QY 355 ATTAATTTAGATTCAGATCAAGGTTTCAAG---GAATACATCCATGCTGTAAACATCTC 411
Db 433 ATCTGTGGACTGGAAGAAGGTTTCTCAGACTACCTTACACAGGCTGCCAACATGTG 492
QY 412 CGTGTGTGTGCTGAGGTGGCTTATTTTATGATGTTCTCATGAAAAAATTTGAATAT 471
Db 493 CGATGATGGGTGCCAGGTAGCTCAGATGATCGACATCTTGTGAAAAACTACAGTAC 552
QY 472 TCCCTCTTAAGTGCATGATTTGGCCACAGCTTGGAGACACACCTGGCTGGGAGCT 531
Db 553 TCGCCTTCCAAAGTCCACCTCATTTGGCCACAGCCTAGGACCCACGTCGGCGGGGAAGCA 612
QY 532 GGTCAAGGATACAGGCTTGGAGAAATAACTGGTGGACCCAGCTGGGCCATTTTTC 591
Db 613 GGAAGTCGATCTCAGGCTTAGGAGAAATTAACAGGATGATCTCTGTAGAAGCAAACTTC 672
QY 592 CACAACTCCAAAGGAGTCAAGCTAGACCCCTCGGATGCCAATTTTGTGAGCTTAT 651
Db 673 GAGGCACTCCTCTGAAGAGTCCGCTTGACCCCTCGGATGCTGACTTTGTGATGTGAT 732
QY 652 CATCAATGAGCTCGCATCTTTGAGCTTGTGTGGAACCATTTGATGCTGTGCT 711
Db 733 CACAGATGAGCTCCCTTGTATCCGCTTGTGGGCTTCGGAACCAACCAATGTGAGG 792
QY 712 CATCTTGACTTTTACCCAAATGGAGGAGCACAATGCCAGATGTGAAGACTTAATTACA 771
```

```

793 CACCTGACCTCTTCCCAACGAGAGAGAGATGCGCGGTGCAAGAA---GATGCT 849
772 CTTTACTGAAATTTAACTCAATGCTTCAAAAGAAATGGCTTCTTCTTCTGACTGT 831
850 CTGTCCAGATTTGACATCGATGATGCTGCTGAGGAGACCGCGGACTTTGTGCTTGT 909
832 AACCATCCCGAAGTTATCAATTTTATGCTGAAAGCAATCTTAACTCTGATCAATTTAT 891
910 AACCATCTGAGAGCTCAAGTACTTCTTGGAGAGATCTTAACTCTGATGCTTCTGCT 969
992 GCTTATCTTTGTAGATCTTACATCTTTTAAAGCAGGAAATGCTTCTTTTGTCCAAA 951
970 GCATACCCCTGTGCTTCTTCAAGGACTTTGAGTCTAAACAAATGCTTCCCTGCGCAGAT 1029
952 GAAGTTGCCCAACAAATGGGTCTATTTTGTGATAGATTTTCACTTCAAAATATGAAGACT 1011
1030 CAAGGCTGCCCAACAGATGGGTCTATGCGGATTAAGTTTCCGGCAAGTCAAGTGA---C 1086
1012 AATGATACATATTTTAAACACAGAGTCCCTTTCCCAATTTGCCCTTTGGAGGCAC 1071
1087 GAGCCACAGAGTTCTTCTTGAACACAGGAGAGCAAGCAAACTTTGACGCTGGAGGTAC 1146
1072 AAATGTCTGTAACTCAGTGGAGCAAGTCACTCAAGCAACTGCTTTCTTCTGCTGA 1131
1147 CGTGTCTTCTGTACTGCTGGAAGAAATGGTCTAC---AGGGCAAGTCAAGTGGCTCTG 1203
1132 GCGCGGCAATTTGGAAACCTGGGAGTTTGGCAATGCTCAGTGGAAACCTTGGAGCAGGC 1191
1204 TTTGGAAGTAGGCAATACACGCCAGTACGATCTTCAAGGGAATTTCAAGCTGCT 1263
1192 ATGACTTACAAATTAATGATGAGATGTTAACTGTTGGAACATTTACAGTGTCTAG 1251
1264 GCTACATTTCCAGTGAAGTTGATGCGCAAGCTCGAGTGGGAACAATTTGAGAAAGTCAAG 1323
1252 TTCACTCGGAAACAACTTTGGAAGATTTCTCAGAAATTAAGTTGGGAGCAGAAATGGTG 1311
1324 TTTCTTGGAAACATCAAGTATTAACCCAGCTTCCCAAGTGGGCGCAGCCAGATC 1383
1312 ATAATACATCTGGAAATATGGAATATTAATCTACCTTCTGTAGCCAAAGATTTATGGA 1371
1384 ACTGTGCAAAAGGAGAGGAGCGGAGGAGTACAACTTCTGTAGTGAAGAGACCGTGGA 1443
1372 CCTATATTTCCAGACCTGAAACCACTGCTAA 1404
1444 GAAGACACTCTGCTCACTCTTCTGCTTTGTGAA 1476

```

```

LT 14
7402
S NITION
Spermophilus tridecemlineatus clone 7G5 pancreatic triacylglyceride
lipase (PTL) mRNA, complete cds.
SSION
ION AF177402
AF177402.1 GI:5762485
ORDS
CE
GANISM
Spermophilus tridecemlineatus (thirteen-lined ground squirrel)
Spermophilus tridecemlineatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Scuriidae; Sciurinae;
Spermophilus.
1 (bases 1 to 2338)
Bauer, V.W. and Andrews, M.T.
Direct Submission
Submitted (13-AUG-1999) Genetics, North Carolina State University,
3513 Gardner Hall, Raleigh, NC 27695, USA
2 (bases 1 to 2338)
Bauer, V.W. and Andrews, M.T.
Direct Submission
Submitted (13-SEP-1999) Genetics, North Carolina State University,
3513 Gardner Hall, Raleigh, NC 27695, USA
MARK
Amino acid sequence updated by submitter
Location/Qualifiers

```

```

source
1..2338
/organism="Spermophilus tridecemlineatus"
/mol_type="mRNA"
/db_xref="taxon:43179"
/clone="7G5"
/tissue_type="white adipose tissue"
1..2338
gene
/feature="PTL"
845..2242
CDS
/feature="PTL"
/codon_start=1
/product="pancreatic triacylglyceride lipase"
/protein_id="AAD51123.2"
/db_xref="GI:5870830"
/translation="MLLVMSLALLGAVAGKEVCYDLRGCFSDSPWSGIVERPELKLVL
PSPADVNTRELLYNQDNYQITADSSRIQSNFKTKRKTPIIHGFIDKEESW
LANMKQFQVESVNCI QDWKGSRTGYTHASONIRIVGAEVAYFVDFLTQIGYSP
SNVTHVHLSHAGAGBAGRTNGAIGRTGLDPAEPCFECTBELRDLPSDAQFVDA
IHTDGAIVPVLNFGMSCTVGHLPFNGGTMPGCKNLSQIVDIDGIVEGTRDFA
ACNHLRSYKTYDTSIVNTPGAFSCASYSVFSANKKCPSPSGGCPOMGHVADRYSGK
TNGVQKFLYNTGDKSNFNRWYKVSFVLSQKVTGCHLLVSLFGNAGNSKQVEYLYKGS
LHPGTHSNEFDSDVDGDLQVRKPIWYNNVINPFLPRVGASSISVERNDGRVAFKFS
AETVEDVILLTLNAC"
BASE COUNT 669 a 535 c 546 g 588 t
ORIGIN
Query Match 24.6%; Score 345.8; DB 10; Length 2338;
Best Local Similarity 57.2%; Pred. No. 3.3e-73;
Matches 793; Conservative 0; Mismatches 567; Indels 27; Gaps 8;
QY 29 TGTTCCTTGGCACATCAAGAGGAAAGAGTTTGTCTATCAAAAGTTAGGGTGTTCCTCAAG 88
DB 870 TGTGCTGGGAGCAGTAGCAGGAAAGAGTCTGCTATGACAGACTCGGCTGCTTCAGTG 929
QY 89 ATGCTTTACATGGACGAGACTTCTTCAACAGAGTTGTTAGGTTTACCTGCTCTCCAG 148
DB 930 ATGACTCCCATGCTCAGGAATCGTGGAAAGACCCCTGAAAGTATTGCCCTGGAGTCCAG 989
QY 149 AGAAGATAACACTCTGTTCTCTACACTATACAAATCCCAATGCTCTATCAGGAGA 208
DB 990 CAGATGTCAACACCCCTCTCTCTGCTACACTTACGAGAACCAAGACAACTATCAACAA 1049
QY 209 TCAGTCGGGTAAATCTTCAACTATCCAAAGCTCTATATTTTGGAAACAGACAGATCACCC 268
DB 1050 TCACTGCGG---ATTTCATCAAGAAATTCAGAGCTCCAAATTTCAAAACAAATAGGAAACCC 1106
QY 269 GTATCAACATAGCTGGT-----GGAAACAGATGGGCAATGGCAGAGACATGTGCA 322
DB 1107 GCTTCATATTACCGGATTCATAGACAGGGAGAGAAAGCTGGCTGGCCACATGTGCA 1166
QY 323 ATGTGTTGCTACAGCTGGAGATATAAATTGCAATTAATTAGATTGGATCAACGGTTTCA 382
DB 1167 AGAAATGTTCCAGGTGGAAAGTGTGAACCTGCTGCTGAGCTGGAGGGCGGCTCCC 1226
QY 383 GG---GAATACATCCATGCTGTAACAACCTCGTGTGTTGTTGCTGAGGTGGCTTATT 439
DB 1227 GGACCGGATACACCCATGCTCCAGAACATCCGGAATCGTGGGTGAGAGTGGCTACT 1286
QY 440 TTATTGATGTTCTCATGAAATAATTTGAATATTCCCTTCTTAAAGTGCACTTGATTTGGCC 499
DB 1287 TTGTTGATTTCCTTCGAGCTCAATTAGGATCTCGCTTCCAAATGCTCATGTCTTGGCC 1346
QY 500 ACAGCTGGGAGACACCTGCTGGGAGAGCTGGTCAAGGATACCAGG---CCTTGGAA 556
DB 1347 ACAGCTGGGTTCCCATGCTGCTGGGAGGCTGGAAAGAGGACCAATGGGGCCCATTTGGAC 1406
QY 557 GAATAACTGGGTTGGACCCAGCTGGGCAATTTTCCACCAACACTCCAAAAGAGTCCAGGC 616
DB 1407 GAATCAGAGATTGGATTCAGCTGAACCTTGTCTTGGAGGTACACCTGAACCTTGTCCGAT 1466
QY 617 TAGACCCCTCGGATGCAACTTTGTTGACGTTTATTAATATCAATGACAGTCCGATCCTCT 576
DB 1467 TGGACCCCGAGTGTCTCAGTTTGTGATGCGATTACACTGTATGTTGCCCAATAGTCC 1526

```

677	TTGAGCTTGGTGTGAACCAATTGATGCTTGTGGTCACTTTGACTTTTACCCAAATGGAG	736
1527	CCAACTTAGGGTTTGGATGAGCCAACTGTGGTCACTAGATTTCTTCCAAATGGAG	1586
737	GGAGCACATGCGAGATGTGAAGACTTAATATACACCTTTTACTGAAATTTAACTTCAATG	796
1587	GAATGAATGCGGGATGTCAAGAATAATT--CTCTCAGATTTGTTGACATAGATG	1643
797	CTTACAAAAGAAATGGCTTCTTCTTTGACTGTAACCATGCGCGAAGTTATCAATTTT	856
1644	GAATTTGGGAAGTACTCGTGAATTTTGAGCATGTAATCACTTAAGAGCTACAAGTATT	1703
857	ATGCTGAAGCAATCTTAATCTGATGCAATTTATGCTTATCTCTTGTAGATCTTACACAT	916
1704	ATACTGATAGTATTGTCAACCCCACTGGCTTTGCTGCAATCTCTTGTGCTCTTATAGT	1763
917	CTTTTAAACGGAATTCCTCTTCTTTTCCAAAGAGGTTGGCCACACATGGGTCAAT	976
1764	TTTCTCTGCAACCAAGTCTTCCCTGTCCAGTGGAGGATGCCACAGATGGGTCAAT	1823
977	TTGCTGATAGATTTCACTTTCAAAAATATGAAGACTAAATGGATCACTATTATTTTAAACA	1036
1824	ATGCTGATAGAT--ATTCTGGAATAACAAATGGAGTGGCCAGAAATTTTATCTGAACA	1880
1037	CAGGTCCTCTTCCCAATTTGGCGTTGGAGGACAAATTTGCTGTGTTAACTCAGTGGAA	1096
1881	CTGGGATAAGAGTAATTTCTCAGCTTGAGGATCAAGGTGTCTGTCACTGTCTGGGC	1940
1097	GCAAGTCACTCAAGGAATGCTTCTTCTGTTGAGCGGGCAATTTGGGAAATCTGGG	1156
1941	AGNAGTTACAGGACATATCTGTTTCTGTTTGGAAATGCA---GGAACTCTAAGC	1997
1157	AGTTTCCCATTTGAGTGAAGAACTTGAGCCAGGATGACATTAACAAAATTAATCGATG	1216
1998	AGTATGAATTTTACAGGGCTCTCTTCTATCCGGGTATACATTTCCAATGAGTTGACT	2057
1217	CAGATCTTAACGTTGGAACATTAACAGGTTTCACTTCTCATCTGGAAACAACTTTGTTG	1276
2058	CTGATGTGAGCTTGGGACTTGGAGGGTCAAAATTTATTTGGTATACAAATGTGATTA	2117
1277	AAGATTTCTCAGAATAAGTTGGGAGCAGAAATGGTGATAATAATACATCTGGGAAATATGAT	1336
2118	ACCATCACTACCCAGAGTGGAGCATCTA---GCATCTCGTGGAAAGAACGATGAA	2174
1337	ATAAATCTACCTTCTGTAGCCAGACATTAATGGGACCTTAATTTCTCCAGAACCTGAAAC	1396
2175	GAGTGTCAAAATTTCTGTAGTGACAGACTGTGAGGGAGGATGTTCTGCTCACCCTCAACG	2234
1397	CATGCTA 1403	
2235	CATGTTA 2241	
T 15		
870		
ITITON	Spermophilus tridecemlineatus (thirteen-lined ground squirrel)	
STON	AF395870 1846 bp mRNA linear ROD 12-JUL-2001	
ON	mRNA, complete cds.	
RDS	AF395870.1 GI:14701914	
TANISM	Spermophilus tridecemlineatus (thirteen-lined ground squirrel)	
E	Spermophilus tridecemlineatus	
ANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;	
	Spermophilus.	
ENCE	1 (bases 1 to 1846)	
HORS	Squire, T.L., Bauer, V.W., Lowe, M.E. and Andrews, M.T.	
LE	Genomic Organization of the Pancreatic Triacylglycerol Lipase Gene	
TRNAL	in a Hibernating Mammal	
ENCE	2 (bases 1 to 1846)	
AUTHORS	Squire, T.L. and Andrews, M.T.	
TITLE	Direct Submission	
JOURNAL	Submitted (28-JUN-2001) Department of Biology, University of	
FEATURES	Minnesota, Duluth, 10 University Drive, Duluth, MN 55812, USA	
source	Location/Qualifiers	
	1. 1846	
	/organism="Spermophilus tridecemlineatus"	
	/mol_type="mRNA"	
	/db_xref="taxon:431179"	
	/tissue_type="pancreas"	
	358..1755	
	/codon_start=1	
	/product="pancreatic triacylglycerol lipase"	
	/protein_id="AAK72259.1"	
	/db_xref="GI:14701915"	
	/translation="MLLWSLALLLGAAGKEVCYDRLGCFSDSPWSGIVERPLKVL	
	PWSADVNTPLFLYTNENODNYQITADSSRIQSSNFKNRKTFRFIHGFIDKGEESW	
	LANNCKMFQVESVNCICVDKMGSGRTGYTQASONTIRIVGAEVAYFVDLRTOLGYS	
	SNHVIGHSLGSHAAAGAGRTWGAIGRTGLDPAEPCEPGLPDLVLPDSDAQFYDA	
	IHTDGAFLVNLGFGMSQTVGHLDFFPENGIEWPGQKNILSIQIVLDIGTWETRFA	
	ACNHLRSYKYYTDSIVNPTFAAPSCASYSVFSANKCTPCPGCGCFMGHAYDIYSGK	
	TNGVQKFLYNTDGSNFSRWYKVSYSQKQVGHILVSLFGNAGNSKQYDIYSGS	
	LHPGYTHSNFDSVDVGDQLQRVKFIWYNNVINPSPRVGASSISVERNDGRVFKFCS	
	AETVREDVLLTLNAC"	
BASE COUNT	509 a 422 c 440 g 474 t 1 others	
ORIGIN		
Query Match	24.5%; Score 344.2; DB 10; Length 1846;	
Best Local Similarity	57.1%; Pred. No. 8.3e-73;	
Matches	792; Conservative 0; Mismatches 568; Indels 27; Gaps 8;	
QY	29 TGTCTTTGGCACATCAAGAGAAAGAGTTTCTCTATGAAGGTTAGGTTGTTTCAAAG	88
Db	383 TCTCTCTGGGAGCAGTAGCAGAAAGAGTCTCTATGACAGCTCGGCTGCTTCAATG	442
QY	89 ATGTTTACCATGACACAGGACTTTCTCAACAGAGTTGGTAGGTTTACCTTGGTCTCAG	148
Db	443 ATGACTCCCATGTTGTCAGGAATCTGGAAGAGACCCCTGAAAGTATTGCTCGAGTCCAG	502
QY	149 AGAATATAACACTGTTTCTCTCTACACTATACAAATCCCAATCCCAATCCATCAGGAGA	208
Db	503 CAGATGTCAACACCGCTTCTCTCTGACACTACAGAGAACCAAGAACATATCAACAA	562
QY	209 TCAGTCCGCTTAATTTCTCAACTATCCAAGCTCATATTTTGGAAACAGACATCACCC	268
Db	563 TCATCGCG---ATTCAACAAGATTCAGAGCTCCAAATTTCAAACAAATAGGAAACCC	619
QY	269 GTATCAACATAGCTGGAT-----GGAAACAGATGGCAATTCGACAGAGACATGTGCA	322
Db	620 GCTTCATTATTTCACGATTCATAGACAGGAGAGAAAGCTGGCTGGCCCAACATGTGCA	679
QY	323 ATGCTGTGTCACAGCTGGAAGATATAAATTCATTAAATTTAGATTGGATCAACGGTTTCAAC	382
Db	680 AGAAATGTTTCCAGGTGGAAGAGTGTGAATCTCTGTGTGACTGGAAGGCGGCTCCC	739
QY	383 GG---GAATACATCATCTGTAAACAATTCCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGT	439
Db	740 GGACCGGATACACCCAGGCTTCCCAAGAACATCCGATCGTGGGTGCAAGTGGCTTACT	799
QY	440 TTATGTAGTGTCTCATGAAATAATTTGATATTTCCCTTCTTAAAGTGCACCTTGTATGCCC	499
Db	800 TTGTTGATTTCTCGGACTCAATTTAGGATGATCTCGCTTTCATGTCCATGTCAITGGCC	859
QY	500 ACAGTGTGGGAGCACACCTGGCTGGGGAAGCTGGGTCAAGGATACCAAGG---CCTTGGAA	556
Db	860 ACAGTGTGGTTCCTCATGAAATAATTTGATATTTCCCTTCTTAAAGTGCACCTTGTATGCCC	919
QY	557 GAATTAACCTGGTGTGAACCGCTGGGCCATTTTTCACCAACACTCCCAAGGAAGTCTAGGC	616
Db	920 GAATCACAGGATTTGATCAGCTGAACCTTCTTGTAGGGGTACACCTGAACTTGTCCGAT	979
QY	617 TAGACCCCTCGGATGCCAACTTTGTTGACGTTATTTCATACAAATGACAGCTCGCATCTCT	676

980 TGGACCCAGTGATGCTCAGTTTGTGGATGCGATTACACCTGATGTTGGCCCAATAGTCC 1039  
677 TTGAGCTTGGTCTTGGACCACTTGATGCTTGTGCTCATCTTGACTTTTACCAAAATGGAG 736  
1040 CCAACTTTAGGGTTTGGATGAGCCAACTGTGGGTCACTAGATTCTTCCAAATGGAG 1099  
737 GGAAGCACATGCCAGGATGTGAAGACTTTAATTACACCTTTACTGAAATTTAACTTCAATG 796  
1100 GAATAGAAATGCCGGATGTGAGAAGATAATT--CTCTCTCAGATTGTTGACATAGATG 1156  
797 CTTACAAAAGAAATGGCTTCTTCTTTGACTGTACCATGCCCGAAGTATCAATTTT 856  
1157 GAATTTGGGAAGTACTCGTACATTTGAGCATGTAACTACCTAAGAAGCTTACAAGTATT 1216  
857 ATGCTGAAAGCAATTCCTAATCCTGATGCAATTTATTGCTTATCCTTTGATGATCCTACACAT 916  
1217 ATACTGATAGTATTGTCACCCCACTGCTTTGTCGATTCTCTTTGGCCTTTATAGTG 1276  
917 CTTTAAAGCAGGAAATTCCTTCTTTTGTTCCTAAGAGGTTGCCCAACAATGGGTCAAT 976  
1277 TTTTCTCTGCAACAAGTCTTCCCTCTCCAAAGTGGAGGATGCCCAAGATGGGTCAAT 1336  
977 TTGCTGATAGATTTCACATTCAAAATATGAAGACTAATGGATCACATTTATTTTAAACA 1036  
1337 ATGCTGATAGAT--ATTCTGGAATAACAATGAGTGGGCCAGAAATTTTATCTGACCA 1393  
1037 CAGGTCCCTTTCCCATTTGCCCTTGGAGGCAAAATGTCTGTGTTAACTCAGTGGAA 1096  
1394 CTGGGATAAAGAGTAATTTCTCACGTTGGAGGTACAAGGTCTGTGCACACTGTCTGGGC 1453  
1097 GCGAAGTCACTCAAGGAATCTGTTCTTCTGTTAGGGGGCAATTTGGGAAACTGGGG 1156  
1454 AGAAGTTACAGGACATATCTGGTTTCTCTGTTTGGAAATGCA--GGAACTCTAAGC 1510  
1157 AGTTTGCCATTTGTCAGTGGAAACTTTGAGCCAGGCACTGATTTACACAAAATTAATCGATG 1216  
1511 AGTATGAAATTTACAAGGCTCTCTTTCATCCGGCTATCTACTATTCCAATGAGTTGACT 1570  
1217 CAGATGTTAAAGTTGGAACATTACAAGTGTTCAGTTTCATCTGGAATAAACATTTGTTG 1276  
1571 CTGATGTGACGTTGGGACTTTCAGAGGGTCAAATTTATTGTTATACAAATGTGATTA 1630  
1277 AAGATTCTCAGAAATAGTTGGAGCAGAAATGGTGATAAATACATCTGGGAAATATGGAT 1336  
1631 ACCCATCACTACCCAGAGTGGGAGCATCTA--GCATCTGGTGGAAAGAACGATGGAA 1687  
1337 ATAAATCTACCTTCTGTAGCCAAAGACATATGGGACCTAATATTCTCCAGAACCTGAAAC 1396  
1688 GAGTGTTCAAATTTCTGTAGTCAGAGACTGTGAGGGGAGTGTCTGCTCACCCTCAACG 1747  
1397 CATGCTA 1403  
1748 CATGTTA 1754

rch completed: January 6, 2004, 10:17:28  
time : 5234 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

cleic - nucleic search, using sw model

n: January 6, 2004, 08:35:07 ; Search time 3190 Seconds  
(without alignments)

10697.017 Million cell updates/sec

US-10-038-517-1

ct score: 1404

ncc: 1 atgcttgaatttgattgt.....agaacctgaaccatgctaa 1404

ng table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

hed: 22781392 seqs, 12152238056 residues

number of hits satisfying chosen parameters: 45562784

um DB seq length: 0

um DB seq length: 2000000000

processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

ase :

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_esti:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estcom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_pbg:\*

27: em\_gss\_vrl:\*

28: gb\_gsa1:\*

29: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Rank	Score	Query Match	Length	ID	Description
1	766.4	54.6	892	10 BG740875	BG740875 602633766
2	510	36.3	1069	11 BC015840	BC015840 Homo sapi
3	447.2	31.9	890	10 BF976776	BF976776 602146609
4	378	26.9	769	10 BG741031	BG741031 602634759

5	373.4	26.6	1528	11	AK028105	AK028105 Mus muscu
6	338.2	24.1	1513	11	BC007025	BC007025 Homo sapi
7	329.8	23.5	682	10	BF246347	BF246347 601854965
8	318.8	22.7	1560	11	AK008021	AK008021 Mus muscu
9	300.6	21.4	757	13	EX259514	EX259514 BX259514
10	286	20.4	778	13	EU416727	EU416727 603671286
11	282.4	20.1	795	13	EU417084	EU417084 603671129
12	277.6	19.8	761	13	EU416864	EU416864 603671065
13	274.8	19.6	780	13	EU416411	EU416411 603551210
14	274	19.5	781	13	EU417156	EU417156 603671487
15	271.8	19.4	741	13	EU416394	EU416394 603670888
16	270.6	19.3	709	13	EU416839	EU416839 603671477
17	269.2	19.2	741	13	EU417450	EU417450 603671653
18	266.8	19.0	689	13	EU417193	EU417193 603671089
19	262.8	18.7	852	12	EU417008	EU417008 603670986
20	261.2	18.6	682	13	EU416310	EU416310 603670586
21	261	18.6	783	13	EU416310	EU416310 603670586
22	260.2	18.5	866	10	BG698664	BG698664 602703028
23	258.8	18.4	821	13	EU489555	EU489555 604128003
24	255.6	18.2	729	13	EU417286	EU417286 603671573
25	254.4	18.1	772	13	EU417314	EU417314 603671534
26	253.6	18.1	677	13	EU417477	EU417477 603671307
27	252.4	18.0	765	13	EU416751	EU416751 603671220
28	252	17.9	717	13	EU417475	EU417475 603551202
29	250.6	17.8	752	13	EU417334	EU417334 603670963
30	247.6	17.6	831	13	EU488059	EU488059 604126109
31	246.8	17.6	839	13	EU592328	EU592328 AGENCOURT
32	244.8	17.4	883	14	CA971009	CA971009 AGENCOURT
33	244.6	17.4	677	13	EU416921	EU416921 603671478
34	243.4	17.3	631	13	EU417237	EU417237 603670805
35	243.4	17.3	700	13	EU491508	EU491508 604131732
36	241.8	17.2	740	10	BG142590	BG142590 ia75h03.Y
37	241.2	17.2	773	13	EU490095	EU490095 604130535
38	239.6	17.1	618	13	EU416910	EU416910 603671484
39	239.6	17.1	869	13	EU593924	EU593924 AGENCOURT
40	238.8	17.0	901	14	CA971112	CA971112 AGENCOURT
41	237.2	16.9	938	13	EU233854	EU233854 AGENCOURT
42	236.8	16.9	682	13	EU416844	EU416844 603670722
43	234.4	16.7	604	13	BQ287503	BQ287503 i166a05.Y
44	233.4	16.6	700	13	EU416315	EU416315 603671118
45	232.8	16.6	944	13	BQ219908	BQ219908 AGENCOURT

#### ALIGNMENTS

RESULT 1  
BG740875  
LOCUS  
DEFINITION  
602633766F1 NCI\_CGAP\_Skn3 Homo sapiens cDNA clone IMAGE:4778825 5',  
mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

892 bp mRNA linear EST 15-MAY-2001  
BG740875.1 GI:14051528  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 892)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
Tissue Procurement: James Cleaver, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA  
Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10635 row: c column: 18  
High quality sequence stop: 849.

JRES  
source

```

Location/Qualifiers
1. .892
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:477825"
/lab_host="DHIOB (T1 phage-resistant)"
/clone_lib="NCI CGAP Skn3"
/site="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."
COUNT 253 a 183 c 196 g 260 t
IN

ary Match 54.6%; Score 766.4; DB 10; Length 892;
at Local Similarity 98.4%; Pred. No. 7.4e-195;
:ches 838; Conservative 0; Mismatches 6; Indels 8; Gaps 6;

486 GCATGTGATGGCCACAGCTTG-GGAGCACACCTGGCTGGGAGCTGGTCAAGGATAC 544
1 GCATGTGATGGCCACAGCTTG-GGAGCACACCTGGCTGGGAGCTGGTCAAGGATAC 60
545 CAGGCTTGGAGAAATAAATGAGTGGTGGACCCAGCTGGCCATTTTCCCAACACTCCAA 604
61 CAGGCTTGGAGAAATAAATGAGTGGTGGACCCAGCTGGCCATTTTCCCAACACTCCAA 120
605 AGGAAGTCAGCTAGACCCCTGGATGCCAATCTTGT-GACGTTATTCATCAAAATGCA 663
121 AGGAAGTCAGCTAGACCCCTGGATGCCAATCTTGTGGAGCTTATTCATCAAAATGCA 180
664 GCTGCATCTCTTTGAGCTTGGTGGAAACCATTCATCTTGTGGTCACTTGTGACTTT 723
181 GCTGCATCTCTTTGAGCTTGGTGGAAACCATTCATCTTGTGGTCACTTGTGACTTT 240
724 TACCAAATGAGGGAAGACACATGCCAGGATGTGAAGCTTAATACACTTTACTGAAA 783
241 TACCAAATGAGGGAAGACACATGCCAGGATGTGAAGCTTAATACACTTTACTGAAA 300
784 TTAACTTCAATGCTTACAAAGAAATGCTCTCTTTTTCACGTGTAACATGCCCA 843
301 TTAACTTCAATGCTTACAAAGAAATGCTCTCTTTTTCACGTGTAACATGCCCA 360
844 AGTTATCAATTTATGCTGAAGCAATCTTAATCCTGATGCATTTATGTTATCCTGT 903
361 AGTTATCAATTTATGCTGAAGCAATCTTAATCCTGATGCATTTATGTTATCCTGT 420
904 AGATCCTACACATCTTTAAAGCAGGAAATGCTCTTTTCCCAAGAGGTTGCCCA 963
421 AGATCCTACACATCTTTAAAGCAGGAAATGCTCTTTTCCCAAGAGGTTGCCCA 480
964 ACAATGGGTCAATTTGCTGATAGATTTTCACTTCAAAATATGAAGCTAATGATCACA 1023
481 ACAATGGGTCAATTTGCTGATAGATTTTCACTTCAAAATATGAAGCTAATGATCACA 540
1024 TAATTTTAAACACAGGTCCTTTCCCATTTCCCGTTGGAGGCAAAATGCTCTGT 1083
541 TAATTTTAAACACAGGTCCTTTCCCATTTCCCGTTGGAGGCAAAATGCTCTGT 600
1084 ARACTCAGTGGAGGCAAGTCACTCAAGAACTGCTCTTTCTGTGTAGCGGGCAATT 1143
601 ARACTCAGTGGAGGCAAGTCACTCAAGAACTGCTCTTTCTGTGTAGCGGGCAATT 660
1144 GGGAAACTGGGAGTTTGGCATTGTCAAGTGAAGAACTTGAGCCAGGCAATGATCACA 1203
661 GGGAAACTGGGAGTTTGGCATTGTCAAGTGAAGAACTTGAGCCAGGCAATGATCACA 720
1204 AAATTAATCGATGCGAGTGTAACTTGGAACTTACAGTGTTCAGTTCATCTGG-- 1260
721 AAATTAATCGATGCGAGTGTAACTTGGAACTTACAGTGTTCAGTTCATCTGGGA 780
1261 -AAAAACAAATTTGTTGAAGATTTCTCAGAAAT-AGTTGGGAGCAGAAAT-GGTGATAAT 1317

```

```

Db 781 AAAAAACCAATTTGCTGTAAGATTCTCAGATAAAGTTGGGAGCAGACATGGGTATACAT 840
QY 1318 ACATCTGGGAAA 1329
Db 841 ACATCTGGGAAA 852

RESULT 2
BC015840 1069 bp mRNA linear HTC 04-MAR-2003
LOCUS Homo sapiens, Similar to pancreatic lipase-related protein 2, clone
DEFINITION IMAGE:4305314, mRNA.
ACCESSION BC015840
VERSION BC015840.1 GI:16198383
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1069)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NTH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcapsb@mail.nih.gov
Tissue Procurement: AFCC/DCRD/DTF
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Paney, Erin Heiton, Mark Kettman, Anuradha
Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 38 Row: c Column: 12
This clone has the following problem: retained intron.

FEATURES
source
1. 1069
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4305314"
/tissue_type="Skin, melanotic melanoma, high MDR."
/clone_lib="NIH_MGC_62"
/lab_host="DHIOB"
/notes="Vector: pDNR-LIB"
BASE COUNT 359 a 170 c 212 g 328 t
ORIGIN

Query Match 36.3%; Score 510; DB 11; Length 1069;
Best Local Similarity 100.0%; Pred. No. 8.8e-126;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 895 TATCCTTGTAGATCTACACATCTTTTAAAGCAGGAAATGCTCTTTTGTCCAAAGAA 954
Db 4 TATCCTTGTAGATCTACACATCTTTTAAAGCAGGAAATGCTCTTTTGTCCAAAGAA 63
QY 955 GTTGTCCCAACATGGTCAATTTTGTGATAGATTTTCACTTCAAAATATGAAGCTAAT 1014
Db 64 GTTGTCCCAACATGGTCAATTTTGTGATAGATTTTCACTTCAAAATATGAAGCTAAT 123
QY 1015 GGATCACAATATTTTAAACACAGGTCCTTTCCCATTTCCCGTTGGAGGCACAAA 1074
Db 124 GGATCACAATATTTTAAACACAGGTCCTTTCCCATTTCCCGTTGGAGGCACAAA 183
QY 1075 TTGCTGTAACTCACTGGAAGCAAGTCACTCAAGGAACGCTCTTCTCTGTGTAGGC 1134

```

184 TTGTCGTGTTAAATCAGTGAAGCGAAGTCACTAAGGAACCTGTTCTTTCTGTTGATGGC 243  
 1135 GGGGCAATGGGAAATCTGGGAGTTTGCCATTGTCAGTGGAAATCTTGAGCCAGGCATG 1194  
 244 GGGGCAATGGGAAATCTGGGAGTTTGCCATTGTCAGTGGAAATCTTGAGCCAGGCATG 303  
 1195 ACTTACACAAAATTAATCATGATCAGATGTTAACTGTTGAAACATTACAGTGTTCAGTTC 1254  
 304 ACTTACACAAAATTAATCATGATCAGATGTTAACTGTTGAAACATTACAGTGTTCAGTTC 363  
 1255 ATCTGGAAGAAATCTGTTTCAAGATCTCAGATAAGTTGGGAGCAGGAATCGTGATA 1314  
 364 ATCTGGAAGAAATCTGTTTCAAGATCTCAGATAAGTTGGGAGCAGGAATCGTGATA 423  
 1315 AATACATCTGGGAAATPATGATATAAATCTACCTTCTGTAGCAGACATTATGGACCT 1374  
 424 AATACATCTGGGAAATPATGATATAAATCTACCTTCTGTAGCAGACATTATGGACCT 483  
 1375 AATATTCTCCAGAACCTGAACCATGCTAA 1404  
 484 AATATTCTCCAGAACCTGAACCATGCTAA 513

T 3  
 776

ITTON  
 2F976776 890 bp mRNA linear EST 22-JAN-2001  
 602146609F1 NIH\_MGC\_62 Homo sapiens cDNA clone IMAGE:4305314 5',  
 mRNA sequence.

SIION  
 ON  
 2F976776

RDS  
 EST.

2F976776.1 GI:12343991

1 (bases 1 to 890)

NIH-MGC

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: ARCC/DCTD/DTP

cDNA Library Preparation: CLONETECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1CM1172 row: b column: 03

High quality sequence stop: 601.

Location/Qualifiers

1. .890

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4305314"

/lab host="NIH10B (T1 phage-resistant)"

/clone lib="NIH MGC 62"

/notes="Organ: skin; Vector: pDNR-LJB (Clontech); Site\_1:

Sfil (ggccgcttcggcc); Site\_2: Sfil (ggccattatggcc);

Double-stranded cDNA was prepared from cell line RNA. 5'

and 3' adaptors were used in cloning as follows: 5'

adaptor sequence: 5'-CAGCGCATATGCGC-3' and 3' adaptor

sequence: 5'-ATTCTAGAGCGCGAGCGGCATG-dT(30)BN-3'

(where B = A, C, G, or T). Average

insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies

contained inserts by PCR. This library was enriched for

full-length clones and was constructed by Clontech

Laboratories (Palo Alto, CA)."

277 a 173 c 200 g 240 t

COUNT

IN

Query Match 31.9%; Score 447.2; DB 10; Length 890;  
 Best Local Similarity 98.6%; Pred. No. 6.4e-109;  
 Matches 493; Conservative 0; Mismatches 3; Indels 4; Gaps 4;

895 TATCTCTGTAGATCTCTACACATCTTTTAAAGCAGGAATGCTTCTTTTCCAAAGAA 954

1 TATCTCTGTAGATCTCTACACATCTTTTAAAGCAGGAATGCTTCTTTTCCAAAGAA 59

955 GGTTCGCCCAACAATGGGTTCATTTTCTGATAGATTTTCACTTCAAAAATATGAAGACTAAT 1014

60 GGTTCGCCCAACAATGGGTTCATTTTCTGATAGATTTTCACTTCAAAAATATGAAGACTAAT 118

1015 GATACACATATTTTAAACACAGGGTCCCTTTCCCATTTGCGCGTGGAGGCACAAA 1074

119 GATACACATATTTTAAACACAGGGTCCCTTTCCCATTTGCGCGTGGAGGCACAAA 178

1075 TTGTCGTGTTAAACTCAGTGGAGGAGTCACTCAAGGAACTGTCTTTCTTCTGTAGGC 1134

179 -TGTCTGTTAAACTCAGTGGAGGAGTCACTCAAGGAACTGTCTTTCTTCTGTAGGC 237

1135 GGGGCAATTTGGGAAACTGGGAGTTTCCCATTTGTCTAGTGGAAACTTGAGCCAGGCATG 1194

238 GGGGCAATTTGGGAAACTGGGAGTTTCCCATTTGTCTAGTGGAAACTTGAGCCAGGCATG 297

1195 ACTTACACAAAATTAATCGATCGATGTTAACTGTTGAAACATTACAGTGTTCAGTTC 1254

298 ACTTACACAAAATTAATCGATCGATGTTAACTGTTGAAACATTACAGTGTTCAGTTC 357

1255 ATCTCGAAAAACATTTCTTTTGAAGATCTCAGATAAGTTGGGAGCAGGAATCGTGATA 1314

358 ATCTCGAAAAACATTTCTTTTGAAGATCTCAGATAAGTTGGGAGCAGGAATCGTGATA 417

1315 ATATCATCTGGGAAATATGGATATAAATCTACCTTCTGTAGCAGACATTATGGACCT 1374

418 ATATCATCTGGGAAATATGGATATAAATCTACCTTCTGTAGCAGACATTATGGACCT 476

1375 AATATTCTCCAGAACCTGAA 1394

477 AATATTCTCCAGAACCTGAA 496

RESULT 4

BG741031

602634759F1 NC1\_CGAP\_Skn3 Homo sapiens cDNA clone IMAGE:4779662 5',

mRNA sequence.

769 bp mRNA linear EST 15-MAY-2001

BG741031

EST.

1 (bases 1 to 769)

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: James Cleaver, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA

Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM10637 row: f column: 15

High quality sequence stop: 766.

Location/Qualifiers

1. .769

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

FEATURES

source

```

/clone="IMAGE:4779662"
/lab host="DH10B (T1 phage-resistant)"
/clone lib="NCI CGAP SKM3"
/notes="Organ: sKin; Vector: pCMV-SPORT6; Site: 1. NotI;
Site: 2. SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."
COUNT 240 a 134 c 164 g 231 t
IN
ery Match 26.9%; Score 178; DB 10; Length 769;
st Local Similarity 99.7%; Pred. No. 2.5e-90;
tches 389; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
1015 GGATCACATTTATTTTAAACACAGAGTCCCTTTCCCATTTGCCGTTGGAGGCACAAA 1074
|||||
1 GGATCACATTTATTTTAAACACACAGAGTCCCTTTCCCATTTGCCGTTGGAGGCACAAA 59
1075 TTGCTGTAACTCAGTGGAGGAGTCTCAAGGAAGTCTTCTTCTGTGTAGGC 1134
|||||
60 TTGCTGTAACTCAGTGGAGGAGTCTCAAGGAAGTCTTCTTCTGTGTAGGC 119
1135 GGGGCAATTTGGGAAACCTGGGAGTCTTCCCATTTGTCAGTGGAAACTTGAGCCAGGCATG 1194
|||||
120 GGGGCAATTTGGGAAACCTGGGAGTCTTCCCATTTGTCAGTGGAAACTTGAGCCAGGCATG 179
1195 ACTTACACAAATTAATCGATGCGAGTGAATTAACCTTGGAAACATTAACAGTGTTCAGTTC 1254
|||||
180 ACTTACACAAATTAATCGATGCGAGTGAATTAACCTTGGAAACATTAACAGTGTTCAGTTC 239
1255 ATCTGGAAAAACATTTCTTTGAAGATCTTCAGATAAATGTTGGAGCAGAAATGGTGATA 1314
|||||
240 ATCTGGAAAAACATTTCTTTGAAGATCTTCAGATAAATGTTGGAGCAGAAATGGTGATA 299
1315 AATACATCTGGAAATATGGATATAAATCTACCTTCTGTAGCCAAAGCATATATGGACCT 1374
|||||
300 AATACATCTGGAAATATGGATATAAATCTACCTTCTGTAGCCAAAGCATATATGGACCT 359
1375 ATATATCTCCAGACCTGAACCATGTCTAA 1404
360 ATATATCTCCAGACCTGAACCATGTCTAA 389

```

```

LT 5
S AK028105 1528 bp mRNA linear HTC 05-DESC-2002
Mus musculus 10 day old male pancreas cDNA, RIKEN full-length
enriched library, clone:1810032A3 product:pancreatic lipase
related protein 1, full insert sequence.
SSION AK028105 GI:26389546
ORDS HTC; CAP trapper.
CE Mus musculus (house mouse)
GANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
THORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TLE High-efficiency full-length cDNA cloning
URNAL Meth. Enzymol. 303, 19-44 (1999)
DLINE 99279253
UNED 10349636
2
THORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
URNAL 20499374
DLINE 11042159
UNED 1
THORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,

```

```

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
Genome Res. 10 (11), 1757-1771 (2000)
11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yanagawa, T.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M.,
Wagner, L., Wasio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bona, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Sessa, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S.,
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1528)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numata, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. 1528
/organisms="Mus musculus"
/mol_type="mRNA"

```

```

/strain="C57BL/6J"
/db_xref="FANTOM_DB:1810032A23"
/db_xref="taxon:10090"
/clone="1810032A23"
/sex="male"
/tissue_type="pancreas"
/clone_lib="PikEN full-length enriched mouse cDNA library"
/dev_stage="10 day old"
70..1491
/notes/unnamed protein product; pancreatic lipase related
protein 1 (MGD) [GI:97723, GB|NM_018874, evidence: BLASTN,
100%, matches=1527]
putative"
/codon_start=1
/protein_id="BAC25750.1"
/db_xref="GI:26389547"
/translation="MLIIWLTIPLELLGAQAQKEVCYDNLGFCSDAPBWAGTAIRPLK
LWPSMCKINTEFLTYNENPTAFOTLQSPSTIEASNFQVAKTRPTIIHGFDKJSE
NYWVDMCKMFOVERVNCICVDMKRGSQITTCQAAANVRVGAQVAMIDILVNFVNY
SASKVHLIGHSIAGVCEAGSRTPGLGRITGLDPVEANFECTPEEVLPDPSDAFVD
VIHTDPAELIPLRGHTGVNGVHDFPFGNGQYMPCKCKNALSQIYVDIDGIVSTRDF
VACNHUHSYKYLIESLNPFDGAYFACSTRDSESNKPCPDQCPQMGHIAQDKFAN
NYSVEPQKFLNTEGAKNPAKRWYRVSLETPSGRTVTKVLSLFGSNGNTRQCDIFRG
I1KPGATNENFADKLVDGTTIEKYKFLNNHVNVPFVKVGAAKITVQKGERTHEFN
CSSETVREDILLTLPLCKSTDM"
1507..1512
/notes="putative"
1528
/notes="putative"
COUNT 398 a 390 c 385 g 355 t
N
ry Match 26.6%; Score 373.4; DB 11; Length 1528;
t Local Similarity 56.5%; Pred. No. 6.4e-99;
ches 799; Conservative 0; Mismatches 596; Indels 18; Gaps 5;
1 ATGCTTGGAAATTTGGATTGTTCATCTTTGTTCTTTGGCACAATCAAGAGGAAAGAGTT 60
|||||
70 ATGCTGATTTCTCTGGACAATCCCTCTTTCTTCTACTGGGAGCAGCTCAGGGGAAAGAGTT 129
|||||
61 TGCATGCAAGGTTAGGCTGTTCAAGATGGTTTACATGGACACAGGACTTTCTCAACA 120
|||||
130 TGCATGACAACCTTGGATGCTTTTCTGATGCTGAGCCCTGGGACAGGACAGTATCAGG 189
|||||
121 GAGTTGGTAGGTTTACCTGGTCTCCAGAGAAGATAAACACTCGTTTCTCGCTCTACACT 180
|||||
190 CCCCTAAAACCTTCTCCCTTGGAGCCCTGAGAAGATCAACACTCGCTTCTCGCTCTACAC 249
|||||
181 ATACACAATCCCAATGCCCTATCAGAGATCAGTCGGGTTAAATTTCTTCAACTATCCAGGCC 240
|||||
250 AATGAGAACCCCACTGCTTTCCAGACTCTTCAGCTTTTCTGACCCATCGACCAATTGAGGCC 309
|||||
241 TCATATTTTGGAAACAGACAAGATCACCCGTTATCAACATAGCTGG-----ATGGAATAACA 294
|||||
310 TCCAAATTTTCAAGTTGCCAGGAGACTCGGTTTCATCATCATGGCTTTCATAGACAAGGA 369
|||||
295 GATGGCAATATGGCAGAGAGACATGTGCAATGTGTGCTACAGCTGGAAGATATAAATGTC 354
|||||
370 GAGGAGAACTGGGTGGTTGCATATGTGCAAGAAACATATTTCCAGTGGAGGAGGTGAACTGC 429
|||||
355 ATTAATTTAGATTGGATCAACGGTTTCACGG---GAATACATCCATCGCTGTAACCAATCTC 411
|||||
430 ATCTGCGTGGACTGGAGAGAGGCTCTCAGACTACCTACACGAGGCTCCCAACANTGTG 489
|||||
412 CGTGTGTTGGTCTGAGGTGGCTTATTTTATTTGATGTTCTCTATGAAAAAATTTGAAATAT 471
|||||
490 CGAGTAGTGGTGCCACGGTGGCTCAGATGATTCACATCTTGTGAGAAATCTTCAACTAC 549
|||||
472 TCCCTTCTTAAGTGCACCTTGATGGCCACAGCTTGGGAGACACCTGCTGGGAGACT 531
|||||
550 TCTGCTTCCAAAGTCCACCTCATTTGGCCACAGCTTGGAGCCCACTGTCAGCGGGAAGCA 609
|||||
532 GGGTCAAGGATACCGGGCCCTTGGAAAGAAATAACTGGGTTGGACCCAGCTGGGCCATTTTTC 591

```

Db	610	GAAGTCGGACTCCAGGCTCTGGCAGGATTTACAGGACTGGATCCTCTAGAGCAAAATTTT	669
Qy	592	CACAACACTCCAAAGGAAGTCAGCTAGACCCCTCGGATGCCCAACTTTGTGTACGTTATT	651
Db	670	GAGGGAACCCCTGAAGGTCGGCTCGACCCCTCGGATGCTGACTTTGTTGATGTCATT	729
Qy	652	CATACAAATGCACTCGCATCTCTTTTGAGCTGTGGTGTGGAAACCAATTGATGCTGTGGT	711
Db	730	CACACAGATGCACTCCCTTGATCCCATTTCTTGGGCTTTGGAAACAAACCAATGTTGGGG	789
Qy	712	CATCTTGACTTTTACCCAAATGAGGGAAGCAGCATGCCAGGATGTGAAGACTTTAATTACA	771
Db	790	CATTTTGACTTCTTCCCAACGAGAGCAGTACATGCCCGGGTGAAGAA--GAAGCT	846
Qy	772	CTTTTACTGAAATTTAACTTCAATGCTTACAAAAAGAAATGGCTTCCTTTCTTTGACTGT	831
Db	847	CTGTCAAGATTTGTGACATCGATGGCATCTGGTCAGGAACCCCGGACTTTGTGGCTTGT	906
Qy	832	AACCATGCCCGAAGTTATCAATTTTATGCTGGAAGCATTTCTTAATCCTGATGCATTTATT	891
Db	907	AACCATCTGAGAAGCTACAGTACTACTTAGAGAGCATCTCTTAACCCAGATGGTTGCT	966
Qy	892	GCTTATCTTTGTAGATCTTACACATCTTTTAAAGCAGGAAATTTGCTTTCTTTCCCAA	951
Db	967	GCATACCCCTGCGCTTCTTACAGGACTTTGAGTCTAACAAATGCTTCCCTGCCAGAT	1026
Qy	952	GAAGTTGCCCAACATGGGTCAATTTTGTGTATAGATTTCATCTTCAAAAATATGAAGACT	1011
Db	1027	CAAGGCTGCCACAGATGGGTCACTATGCTGATAAGTTT---GCCAACCAATACAAGTGT	1083
Qy	1012	AATGGATCACATTTATTTTAAACACAGGCTCCCTTCCCATTTGCCCCGTTGGAGGCAC	1071
Db	1084	GAGCCACAGAAGTTCTTCTGAACAAGGAGCAAGAACCTTTGCAAGCTGGAGTAC	1143
Qy	1072	AAATGCTGTGTTAACTCAGTGGAAAGCAAGTCACTCAAGGAACCTGCTTTCTGTGTGA	1131
Db	1144	CGTGTTTCTTGTGATCTCTGGAAGACGGTCAAC---AGGCCAAGTCAAGTGTCTCTG	1200
Qy	1132	GGCGGGCAATTCGGAAACTGGGGAGTTTGCCATTGTCAGTGGAAAACTTGAGCCAGGC	1191
Db	1201	TTTGGAAAGTAAACGGCAATCTGCCAGTGTGATATCTTCAGGGAAATTAATCAAGCCGGT	1260
Qy	1192	ATGACTTACACAAAATTAATCGATGCAGATGTTAAAGTTGGAAACATTAACAAGTTTCAG	1251
Db	1261	GCTACGCATTCCAAATGAGTTTGATGCCAAGCTGATGTGGAAACATTCAGAAAGTCAAG	1320
Qy	1252	TTTCATCTGGAAAAACATTTGTTGAGATTCACAGATAAGTTGGAGCAGGAAATGGTG	1311
Db	1321	TTTCTTTGGAATAACACAGTGGTAAACCCCAAGTTTCCCAAGTGGGGCAGCAAGATC	1380
Qy	1312	ATAATACATCTCGGAAATATGGATATAAATCTACCTCTGTAGCCCAAGACATTAATGGGA	1371
Db	1381	ACTGTCAAAAGGGGAGGAGCGGACAGAGCAACATCTGTGTAGTGAAGAGACCGTGAGA	1440
Qy	1372	CCTAATAATCTCAGAACTGAACCATGCTAA	1404
Db	1441	GAAGACATCTGCTCACTCTCTTGGCTTGTAA	1473

RESULT 6

BC007025

LOCUS BC007025

DEFINITION Homo sapiens, pancreatic lipase, clone IMAGE:13950244, mRNA.

ACCESSION BC007025

VERSION BC007025.1

KEYWORDS GI:14711980

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1513)

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. Strausberg, R.

RESULT 6  
 BC007025  
 LOCUS  
 DEFINITION Homo sapiens, pancreatic lipase, clone IMAGE:3950244, mRNA.  
 ACCESSION  
 BC007025  
 VERSION  
 BC007025.1 GI:14711980  
 KEYWORDS  
 HTC.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 1513)  
 AUTHORS  
 Strausberg, R.

## TLE

Submitted (30-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

## MARK ENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: CLONTECH

CDNA Library Preparation: CLONTECH Laboratories, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 16 Row: i Column: 8  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 339596  
This clone has the following problem: frame shifted.

## UBES source

Location/Qualifiers  
1. .1513  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="LocusID:5406"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3950244"  
/tissue\_type="Pancreas"  
/clone\_lib="NIH\_MGC\_78"  
/lab\_host="DH10B"  
/note="Vector: pDNR-LIB"  
455 a 320 c 357 g 381 t

## very Match

24.1%; Score 338.2; DB 11; Length 1513;

## st Local Similarity

56.8%; Pred. No. 1.9e-79;

## atches 787; Conservative

0; Mismatches 573; Indels 25; Gaps 8;

29 TGTCTTCGGCATCAAGAGAAAGATTGCTTGAAGGTTAGGTTGTTTCAAAG 88  
52 TGTCTTCGGAGCAGTAGCAGGAAAGAGTTTGTCTACGAAGACTCGGCTTCAGTG 111  
89 ATGTTTACATGGACAGGACTTCTTCAACAGAGTTGGTAGGTTTACCCCTGGTCTCCAG 148  
112 ATGACTCCCATGGTTCAGGAATTACGGAAGACCCCTCCATATATTGCTTGGTCTCAA 171  
149 AGAAGATAAAGACTGTTTCTGCTCTACACTATACACAATCCCAATGCTTATCAGGAGA 208  
172 AAGATGTCAACACCCGCTTCTCTTATATCTATGAGAACCAACAATTCAGAGAG 231  
209 TCAGTCGGTTAATTTCTCAACTATCAAGGCTCATATTTTGGAAACAGACATCACCC 268  
232 T---TGCCGAGATTTCATCAAGCATCAGTGGTCCAAATTTCAAACAATAGAAAACTC 288  
269 GTATCAACATAGCTGGAT-----GGAAACAGATGGCAATTCGAGAGACATGTGCA 322  
289 GCTTTATTATCATGATTTCATAGACAGGGAGAGAAAACTGGCTGGCCAAATGTGCA 348  
323 ATGTGTGTGTACAGCTGGAGAGATATAAATTTGATTAATTTAGATTGGATCAACGGTTTCA 382  
349 AGAATCTGTTCAAGTGGAAAGTGTGAATGATCTGTGTGGACTGGAAAGTGGCTCCC 408  
383 GGAAT-AGATCATCTGTAAACAATCTCCGTGTGTGTGTGTGTGTGTGTGTGTGTGT 441  
409 GAATCGATACACAGACCTCGCAGAACATCAGGATCGTGGGAGCAGAGTGGCATATTT 468  
442 ATTGATGTTCTCATGAAAAAATTTGAATATTCCTTCTAAAGTGCACTTATTGGCCAC 501  
469 GTTGAATTTCTTCAGTCGGCGTTTCGGTTACTACCTTCCACGTGCAATGATGGCCAC 528

QY 502 AGCTTGGGAGCACACCTGCTGGGGAGCTGGGTCAAGGATACCAGG---CCTTGGAGA 558  
Db 529 AGCTGGTGGCCACACCTGCTGGGGAGGCTGGAGGAGAACCAATGGGACCATGGACGC 588  
QY 559 ATAACCTGGGTTGGACCCAGCTGGGGCCATTTTTCACACACATCCCAAGGAAGTCAGGCTA 518  
Db 589 ATCAGAGGTTGGACCCAGCAGAACCTTCTTTCAGGGGACACCTGAATTTAGTCCGATTG 648  
QY 619 GACCCCTCGGATGCCAACTTTTGTGACGTTATTATACAAATGACAGCTCGCATCTCTTT 678  
Db 649 GACCCAGCATGCCAAATTTTGTGATGTAATTCACAGATGGTGGCCCATAGTCCCC 708  
QY 679 GAGCTTGGTGTGGAAACCAATGTAGTGTGTGTCATTTGACATTTTACCAAATGAGGAG 738  
Db 709 AATTTGGGTTTGGAAATGAGCAAGTCGTGGGCCACTAGATTCTTTCCAAATGAGGA 768  
QY 739 AAGCAGATGCCAGTGTGAGACTTTAATTACACCTTTTACTGAAATTTAACTTTCAATGCT 798  
Db 769 GTGAAATGCTGTGATGTAAAGAAACAAT---CTCTCTCAGATTTGGACATAGACGNA 825  
QY 799 TACAAAAAGAAATGGCTTCTTCTTTGACTGTAAACCATGCCGGAAGTTTATCAATTTAT 858  
Db 826 ATCTGGGAGGACTCGAGACTTTGGGCTGTAACTAATTAAGAGCTACAAATATTAC 885  
QY 859 GCTGAAAGCAATCTTAATCTGTATGCAATTTATGCTTATCTTGTAGATCCTACATCT 918  
Db 886 ACTGATAGCATCGTCAACCCCTGTGCTTGTGGATTCCCTGTGCCCTTCTTACAACTGTC 945  
QY 919 TTTAAAGCAGGAAATGCTTCTTCTTTTGTCCAAAGAGGTTGCCCAACAATGGGTCAATTT 978  
Db 946 TTCACTGCAACAGTGTTCCTTGTCCAGTGGAGGCTGCCACAGATGGTCACTAT 1005  
QY 979 GCTGATAGATTCATCTTCAAAAATATGAAGACTAATGAGTACATTTATTTTAAACACA 1038  
Db 1006 GCTGATAGATATCTCTGGGAAA---ACAAATGATGTGGGCCGAGAAATTTTATCTAGACAT 1062  
QY 1039 GGGTCCCTTTCCCATTTGCCGTTGGAGGCACAAATTTGTCTGTAAACTCAGTGGAGC 1098  
Db 1063 GGTATGCCAGTAATTTTGCAGTTGGAGGTATAGGTATCTGTACATCTGTGGAAA 1122  
QY 1099 GAAGTCTCAAGGAATCTTCTTCTTGTGTAGGCGGGCAATTTGGGAAAACCTGGGGAG 1158  
Db 1123 AAGGTAC---AGGACACATACTAGTTCTTCTTGTGGAAATAAAGGAAACTCTAAGCAG 1179  
QY 1159 TTTGCCATGTCACTGGAAACACTTGAGCCAGGCACTGACTTTACACAAAATTAATTCGATCA 1218  
Db 1180 TATGAAATTTCAAGGGCACTCTCAAAACAGATAGTACTCAATTCGAATTTGACTCA 1239  
QY 1219 GATGTTAACTGTGGAAACATTTACAAGTGTTCAGTTTCACTCTGGAAAAAACATTTGTTGAA 1278  
Db 1240 GATGTGATGTTGGGGACTTGCAGATGGTTAAATTTATTTGGTATTAACAATGTGATCAAC 1299  
QY 1279 GATTTCTCAGATAAGTTGGGAGCAGAAATGGTGTAAATATACATCTGGGAATATGATAT 1338  
Db 1300 CCAACTTTTACCTAGAGTGGGAGCATCCAGATATATAGTGGAGAC---AAATGTTGGAAA 1356  
QY 1339 AAATCTACTCTTCTTAGCCAGACATTTATGGGACCTTAATTTCTCCAGAACCTTGAACCA 1398  
Db 1357 CAGTTCACTTCTGTATCCAGAAACCGTCCAGGAGGAGAGTTCTGTCTACCTTCACACCG 1416  
QY 1399 TGCTA 1403  
Db 1417 TGTTA 1421

## RESULT 7

BF246347

## LOCUS

601854965F1 NIH\_MGC\_57 Homo sapiens cDNA clone IMAGE:4074544 5',

## DEFINITION

mRNA sequence.

## ACCESSION

BF246347

## VERSION

BF246347.1 GI:11160614

BF246347 682 bp mRNA linear EST 14-NOV-2000  
601854965F1 NIH\_MGC\_57 Homo sapiens cDNA clone IMAGE:4074544 5',  
mRNA sequence.  
BF246347  
BF246347.1 GI:11160614

DRDS EST. Hom sapiens (human)  
 DE Hom sapiens  
 JANI SM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 682)  
 RENCE NIH-MGC http://mgi.nci.nih.gov/  
 PHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 FLE Unpublished  
 JRNAL Contact: Robert Strausberg, Ph.D.  
 ENT Email: cga@nci.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: CLONTECH Laboratories, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LICM925 row: b column: 17  
 High quality sequence stop: 552.  
 Location/Qualifiers  
 1. 682  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4074544"  
 /tissue\_type="glioblastoma"  
 /lab\_hosts="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH MGC 57"  
 /notes="Organ: Brain; Vector: pDNR-LIB (Clontech); Site\_1:  
 SfiI (ggcgccctcgcc); Site\_2: SfiI (ggccattatggc);  
 Double-stranded cDNA was prepared from cell line RNA. 5'  
 and 3' adaptors were used in cloning as follows: 5'  
 adaptor sequence: 5'-ATTCAGCGCCATTATGCC-3' and 3' adaptor  
 sequence: 5'-ATTCAGCGCCATTATGCC-3' (30)BN-3'  
 (where B = A, C, G, or T). Average  
 insert size 1.55 kb (range 0.9-4.0 kb). 12/15 colonies  
 contained inserts by PCR. This library was enriched for  
 full-length clones and was constructed by Clontech  
 Laboratories (Palo Alto, CA)."  
 COUNT 206 a 130 c 155 g 190 t 1 others  
 IN  
 ary Match 23.5%; Score 329.8; DB 10; Length 682;  
 at Local Similarity 95.8%; Pred. No. 2.2e-77;  
 :ches 392; Conservative 0; Mismatches 12; Indels 5; Gaps 5;  
 1001 ATATGAAGACTAATGGATCACAATATTTTAAACACAGGGTCCCTTTCCCAATT-TGCC 1059  
 1 ATATGAAGACTAATGGATCACAATATTTTAAACACAGGGTCCCTTTCCCAATTGTGCC 60  
 1060 CGTTGGAGGCACAATTTGCTGTAACTCAGTGGAGCGAGTCACTCAAGAACTGTC 1119  
 61 CGTGGAGGCACAATTTGCTGTAACTCAGTGGAGCGAGTCACTCAAGAACTGTC 120  
 1120 TTTCTTCTGCTAGCGGGGCAAT-TGGGAAACTGGGGAGTTTGCCA-TTGTCACTGGAA 1177  
 121 TTTCTTCTGCTAGCGGGGCAATTTGGGAAACTGGGGAGTTTGCCAATTGTCACTGGAA 180  
 1178 AACTTGACCGAGGATGACTACACAAATTAATCGATGAGATGTTACGTTGGAACA 1237  
 181 AACTTGACCGAGGATGACTACACAAATTAATCGATGAGATGTTACGTTGGAACA 240  
 1238 TTCAAGTGTTCAGTTCATCTGGAAGAAACATTTGTTGAAGATTCCTCAGATAAGTTGG 1297  
 241 TTCAAGTGTTCAGTTCATCTGGAAGAAACATTTGTTGAAGATTCCTCAGATAAGTTGG 300  
 1298 GAGCAGAAAT-GGTGATAATACATCT-GGGAATATGGATATAATCTACCTTCGTAG 1355  
 301 GAGCAGAAATGGTGATAATACATCTGGGAAATATGGATATAATCTACCTTCGTAG 360  
 1356 CCAAGACATTATGGACCTATATTTCTCCAGAACTGAAACCATGCTAA 1404

Db 361 CCAAGACATTATGGACCTAATATCTCCAGAACCTTGGAAACCATGCTAA 409  
 RESULT 8  
 AK008021 LOCUS  
 DEFINITION Mus musculus adult male small intestine cDNA, RIKEN full-length  
 enriched library, clone:201001015 product:pancreatic  
 lipase-related protein 2, full insert sequence.  
 AK008021  
 ACCSSION  
 VERSION AK008021.1 GI:12841949  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 99279253  
 10349636  
 REFERENCE 2  
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 20499374  
 REFERENCE 3  
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
 Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, K.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 20530913  
 11076861  
 REFERENCE 4  
 AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
 Arakawa, T., Harada, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yananaka, I.,  
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,  
 Quackenbush, J., Schirml, L. M., Staubli, F., Suzuki, R., Tomita, M.,  
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
 Carninci, P., De Bona, M. P., Brownstein, M. J., Buit, C.,  
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
 Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
 Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,  
 Wyshak-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S.  
 and Hayashizaki, Y.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 21085660  
 11217851  
 REFERENCE 5  
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research  
 Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)



RENCE  
THORS

6 (bases 1 to 1560)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanganaki, T., Hara, A., Hayatsu, S., Hiramoto, K., Hirokawa, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasekawa, T., Kato, H., Kawaji, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

ENT  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Science Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGCCAGGACGCTTTTCTTTTCTTN 3'], cDNA was prepared by using thermostable thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 20.0. Second strand cDNA was prepared with the primer adapter of sequence[5'...

GAGAGAGAGATTCTCGAGTTAATTAATAATCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end; XhoI; 3' end; SstI.

URES  
source

Location/Qualifiers

```
1. .1560
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM,DB:2010001015"
/db_xref="MGI:1905230"
/db_xref="taxon:10090"
/clone="2010001015"
/sex="male"
/tissue_type="small intestine"
/clone_lib="RIKEN full-length e
/clone_tag="s4ul"
```

misc feature

114. 1318  
/note="pancreatic lipase-related protein 2  
(MGD|MG1:1336202)

```

/db_xref="MGI:1336202"

```

polyA\_signal 1543. 1548

```

...
/note="putative"
...

```

```
polyA_site  
1560  
/notc=//m...atg.../
```

405 a	398 c	391 a	366 f
/note=="putative"			
COUNT			

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

ery Match 22.7%; Score 318.8; DB 11; Length 1560;

st Local Similarity	55.0%;	Pred. No. 3.2e-74;
---------------------	--------	--------------------

tches 760; Conservative 0; Mismatches 602; Indels 20; Gaps 6;

35 TTGGGCA CATCAGCGCAAGAGAGTTTCCTATCATACCTTACCTCTTCAATGATCGTT 94

[illegible]

145 TGGCCACAGTCGGAGGGAAGAGGTTTGCTATGGACACCTTGGATGCTTTCCAATGACA 204

95 TACCATGGACCAGGACTTCTCAACAGAGTTGGTAGGTTACCCCTCGTCTCCAGAGAAGA 154

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

205 AGCCCTGGGCAGGGATGATTCAGCGGCCTTCGAAGATATTCCTTGTCCTCCAGAGGACA 264

155 TAAACACTCGTTTCTGTCTCTACATATACCAATCCCAATGCCATTCAGAGATCAAGT 214  
265 TTGATACCCGCTTTCTTCTGTACACAAACGAGAATCCAAAACAACTACCAATAATCTCTG 324  
215 CGGTAAATCTTCAACTATCCAAAGCCTCATATTTTGGAAACAGACAAAGATCACCGCTATCA 274  
325 CCACGACCCAGCCACCATCAATGCTTCCAACTCGACCGCAAGACACGTTTCA 384  
275 ACATAGCTGGATGGAAAAACAGATGGCAAA-----TGGCAGAGAGACATGTGCAATGTGT 328  
385 TTATCCATGGAATTCATAGACAAGGGAGAAAGAGCTGTCTCTGGACATGTGCAAGAAAA 444  
329 TGCTACAGCTGGGAAGATATAAAATTGCATTAAATTTAGATTGGATCAACGGTTTACGG--G 385  
445 TGTTTCAAGTGGAGAAAGTGAATCTGATTTTGGTTTGACTGGAAGCGGGATCCAGGACTG 504  
386 AATACATCCATGCTGTAAACAATCTCGTGTTGTGTGGCTGAGGTGGCTATTATTATTTG 445  
505 AATACACCAAGCCTCCTCAACACTCGGCTCGTGGGCGCTGAGATTGCCCTTTTGTGTGC 564  
446 ATGTTCTTCATGAAAAAATTTGAAATATCCCTTCTTAAAGTGCACTTGATTTGGCCACAGCT 505  
565 AAGTGTGTCGACAGAGATGGGTACAGCCACAGAAACGTGCACCTCATCGGCCACAGCT 624  
506 TGGAGACACACTTGGCTGGGAGCTGGGTCAAGATACAGGCC---TTGGAGAGATAA 562  
625 TGGGCTCACATGTGTGTGGGAGACAGGACAGAGGCTGGAGGGCCATGTGGGCAAGHATCA 684  
563 CTGGGTTGGACCCAGCTGGGGCCATTTTTCACAAACACTCCAAAGGAAGTCAGGCTAGACC 622  
685 CAGGGCTGGACCTGCAGAGCCCTGTTTCCAAGGCTGCGCGAGGAAGTTCGGTTGGACC 744  
623 CCTCGGATGCCACTTGTGTGAGTTATTCATACAAATGCAGCTCGCATCCTCTTTTGAGC 682  
745 CATCGGATGCCATGTTTGTAGACGTATCCACACTGACAGTGCSCCCCATTAATCCCTTATC 804  
683 TTGTTGTGGAACCATTGATGCTGTGTGTGTCATCTTGACTTTTACCAAATGGAGGGGAAGC 742  
805 TAGTGTTCGGAATGAGCCAAAAGGCTGCGCATCTGGAATTTCTTTCCCAACGGAGGGAAG 864  
743 ACATGCCAGAGATGGAAGACTTAATACACCTTTACTGAAATTTAACTTCAATGCTTTACA 802  
865 AATTCGCGGAGATGTAGAA--AAACATCCTTCCACATTTGTTGATATCAATGGAATAT 921  
803 AAAAAGAAATGGCTTCCTTTCTTTGACTGTAACCATGCCCCGAAGTTATCAATTTATGCTG 862  
922 GGGAAAGAACTCGAAACTTCGACGCTGCAACCACTACGAGGCTACAAGTACTATGCGCA 981  
863 AAGCATTTCTTAATCTGTATGATTAATTTGCTTTATCTCTGTAGATCCTACACATCTTTTA 922  
982 GCAGATCTCTCAACCTGACGGC--TTCTGGGTACCCCTGTCTCTCTATGAGAGTTTC 1039  
923 AAGCAGGAAATTCCTCTTTTGTTCAAAGAGGTTGCCAACCAATGGGTCAATTTTGCTG 982  
1040 AGCACAATGACTGTTCCCTTGTTCAGAACACGGGCTGTCCTTCCAAATGGGCACTATGCTG 1099  
983 ATAGATTTCACTTCAAATAATACAGACTTAATGGATACATATTTTAAAAACACAGGGT 1042  
1100 ATCAATTTGAGGGAAACT---GCTACTGTGGAAACAACCTTTTCTGAAATACAGGAG 1156  
1043 CCCTTTCCCAATTTGCCGTTGGAGGCAAAATGTCTGTGTTAAACTCAGTGGAGCGAAG 1102  
1157 ACAGTGTGTAATTTATCTGTTTGGAGATACAAAGTATCAGTCACTATCAGGAGCAAAAGA 1216  
1103 TCATCAAGGAACGTGCTTTCTTCTGTGTAGGCGGGCAATTTGGGAAAACTGGGAGTTTG 1162  
1217 AGCTGAGTGGGTACATCTTGTGTGCTCTGTACGGATGTACGGAATCTCAAGACGATG 1276  
1163 CCATTTGTCAGTGGAAAACTTGACCGACGCGATGACTTACAAAAATTAATCGATCCAGATG 1322  
1277 AAGTTTTCAAGGATCCCTTCCAAACAGAGCAAGATATATTCTGTGAATTCATGATGTGACG 1336

1223 TTACGTTGGAACATACAAAGTGTTCAGTTCATCTCGGAAAAAACAATTTGTTGAAGATT 1282  
 1337 TCAACGTTGGAGAAATCCAGAGGTTAAAGTTCCTCTGGAACAACAAAGGTGATAATCTCT 1396  
 1283 CTGAGATAAAGTTGGGAGCAGAAATGGTGATAAATACATCTCGGAAATATGATGATAAAT 1342  
 1397 TCGGGCCCATATGGGAGCTTCACAAATCACAGTGAACGCGGTAAAGCGGAAGAGT 1456  
 1343 CTACCTTCTGTAGCCAGACATTTATGGGACCTTAATATTTCCAGAACCTCGAAACCATGCT 1402  
 1457 TCAACTTCTGTAGCAGCAACACTGTGCATGAAGATGTTGCGAGTCTCTGTACCCCTTGT 1516  
 1403 AA 1404  
 1517 AA 1518

IT 9  
 514

ITITION BX259514 757 bp mRNA linear EST 27-FEB-2003  
 BX259514 AGENAE Gallus multi-tissues normalized library  
 (gcag) Gallus gallus cdna clone scag0007c.o.07 5prim, mRNA  
 sequence.

STION ON BX259514 GI:28582112  
 RDS EST.

ANISM Gallus gallus (chicken)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.

1 (bases 1 to 757)  
 Herault, P., Le Meuth-Metzinger, V., Desert, C., Retout, E., Piumi, P.,  
 Klopp, C. and Douaire, M.  
 Construction and primary characterization of chicken normalized  
 multi-tissue cDNA libraries

Unpublished  
 Contact: Douaire M  
 INRA, UMR INRA-ENSAR Genetique Animale  
 65, rue de Saint-Brieuc, RENNES cedex, 35042, FRANCE

Tel: +33 (0) 2 23 48 54 63

Fax: +33 (0) 2 23 48 54 70

Email: Madeleine.Douaire@roazhon.inra.fr

Sequence cleaned of vector, adaptor and repetitions. Contact us  
 at signasupport@jouy.inra.fr to obtain the chromatogram of this  
 sequence.

Plate: 0007 row: 0 column: 7

Seq primer: M13R.

Location/Qualifiers

1. /757

/organism="Gallus gallus"

/mol\_type="mRNA"

/db\_xref="taxon:9031"

/clone="scag0007c.o.07"

/tissue\_type="multi-tissues"

/dev\_stage="from embryos to adults"

/lab\_host="DH10B"

/clone\_lib="AGENAE Gallus gallus multi-tissues normalized

library (gcag)"

/notes="vector: pT73D-pac; tissues: brain, embryos, kidney  
 , multi-tissues, muscle, pancreas, skin, testis, liver,  
 adipose tissue, granulosa, utero-vaginal gland, oviduct,  
 small follicle, ovary, hypothalamus, pituitary gland,  
 ileon, jejunum, caecum, duodenum, spleen, fabricius gland  
 , bone marrow, thymus, hematopoietic progenitor cells.

Clone distribution : AGENAE Resource centre. Francois  
 PIUMI, Francois, Piumi.inra.fr, INRA, CEA Radiobiologie et  
 Etude du genome (LREG), Domaine de Vilvert, 78352,  
 Jouy-en-Josas cedex, FRANCE"

223 a 156 c 191 g 186 t 1 others

COUNT  
 N

ry Match

21.4%; Score 300.6; DB 13; Length 757;

Best Local Similarity 64.6%; Pred. No. 1.7e-69;  
 Matches 482; Conservative 0; Mismatches 255; Indels 9; Gaps 2;

QY	22	GCATTCCTGCTTTGGGCACATCAAGAGGAAAAAGTGTGCTATGAAGGTTAGGGTGT	81
Db	10	GCATATTTCTTCTGACACAGAGGCTCTGAAGTTTGTATGACAGGCTGGGATGT	69
QY	82	TTCAAAAGATGGTTTACCATGGACAGGACTTTTCTCAACAGAGTTGGTAGTTTACCCGTGG	141
Db	70	TTCAAGATGATATGCCATGGTCTGAACTGCAGAGAGCAAAATTTATAGGTTTACCCGTGG	129
QY	142	TCTCCAGAGAGATAAACACATCGTTTCCTGCTCTACACTATACACAAATCCCAATGCCCTAT	201
Db	130	AGCCAGAGAGATAGGTACCCAGTTCTCTTACACAGAGAAATGGTAAATAGCAAT	189
QY	202	CAGGAGATCAGTGGGTTAAATCTTCAACTATCCAAAGCCTCATATTTTGGAAACAGACAAG	261
Db	190	CAGGAGATATCTGCAGTTAATCCCTCAACAATTTGGGAGCTCAAACTTCAAAACAAGCAGG	249
QY	262	ATCACCCGTATCAACATAGCTGGATGGAAACAGATGSCAAA-----TGCAGAGAGAC	315
Db	250	AAAGACAAGATTTGTTGTCATGGATTTATAGTGAAGAGAGAGGCTGGACCTCAGAT	309
QY	316	ATGTGCAATGTGTTGCTACAGCTGGAAGATATAAATTCATTAAATTTAGATTGGATCAAC	375
Db	310	CTGTGCAAGAGGATGCTTACAGTGGAGATGTGAATGCTATTGCTGTAGACTGGAAAAA	369
QY	376	GGTTACGG---GAATACATCATGCTGTAAACAAATCTCCGTGTTGTTGGTGTGAGGTG	432
Db	370	GGTCAAGGTTGTCAGTACTCTCAGGCTTCAAAACAACGTCGCTGTTGTTAGTGTCTCAAAAT	429
QY	433	GCTTATTTTATGATGTTCTCATGAAAAAATTTGAATATCCCTTCTTAAAGTGCACCTTG	492
Db	430	GAATATTTTATCAGGCTGTGCGGATCAATATTTACTCTTTCAGCAATGTCCACATC	489
QY	493	ATTGCCACAGCTTCGGAGCACACCTGCTGGGAAGCTGGGTCAAGGATACCAAGGCTT	552
Db	490	ATTGTCACAGCTTTGGAGCACACGTCGAGGAGGCTGGCAAGAGAGGAGGAGGCTT	549
QY	553	GGAGAAATACCTGGGTTGGAGCCAGCTGGGCAATTTTCCACACACTCCCAAGAGAGTC	612
Db	550	GGAGAAATACCTGGGACTGGACCTGCCCAACCTTATTTTCAAGACACTCCAAATGAAAGTC	609
QY	613	AGGCTAGACCCCTCGGATGCCAACTTTTGTGACGTTATTTCATAAATGACGCTCGCATC	672
Db	610	AGGCTGATAAAGCGATGCAAGGTTTGTGATGTTATCCACACAGATACAGCTCCCAT	669
QY	673	CTCTTTGAGCTTGGTTGGAAACCATGATGCTTGTGGTCACTTTGACTTTTACCCAAAT	732
Db	670	ATTCGCCACTTAGGTTTTTGGAAATGGCCCAAGCAATAGGACATCTTGACTTTTATCCCAAC	729
QY	733	GGAGGAGACACATGCCAGGATGTGA	758
Db	730	GGAGAGTGGAAATGCCANGATGTGA	755

RESULT 10

BU416727

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Gallus gallus (chicken)

Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 778)

Boardman, P. E., Sanz-Ezquerro, J., Overton, I. M., Burt, D. W., Bosch, E.,

Fong, W. T., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.

A Comprehensive Collection of Chicken cDNAs

TITLE

BU416727 778 bp mRNA linear EST 29-NOV-2002  
 603671286F1 CSEQRBL07 Gallus gallus cdna clone CHES1615121 5', mRNA  
 sequence.

BU416727

BU416727.1 GI:25909398

EST.

Gallus gallus (chicken)

Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 778)

Boardman, P. E., Sanz-Ezquerro, J., Overton, I. M., Burt, D. W., Bosch, E.,

Fong, W. T., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.

A Comprehensive Collection of Chicken cDNAs

UNRAL  
LINE  
UBMED  
ENT

Curr. Biol. 12 (22), 1965-1969 (2002)  
22335534  
12445392

Contact: Simon Hubbard  
Department of Biomolecular Sciences  
University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK  
Tel: 01612008930  
Fax: 01612360409  
Email: Simon.Hubbard@umist.ac.uk.

URES  
source

1. .778  
/organism="Gallus gallus"  
/mol\_type="RNA"  
/strain="Layer and broiler"  
/db\_xref="taxon:9031"  
/clone="CHEST615121"  
/sex="Male and female"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="CSEORBL07"  
/note="Organ: pancreas; Vector: pBluescript II KS(+);  
Site 1: EcoRI; Site 2: NotI; Modification of pBluescript  
II KS(+) [Stratagene] vector to accommodate cDNA produced  
with the T-trimmed protocol (Construction of  
uni-directionally cloned cDNA libraries from messenger RNA  
for improved 3' end DNA sequencing by Glenn Fu, et al.  
U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with  
NotI and EcoRI. Ligate in double stranded adaptor  
containing BspI and BamHI sites  
[5'ggccgcgtgcagcccgatccgagaaag]  
[5'aattcttttttcggatccgggtgcagc]"

: COUNT 233 a 159 c 194 g 191 t 1 others

IN

ery Match 20.4%; Score 286; DB 13; Length 778;

st Local Similarity 64.4%; Pred. No. 1.5e-65;  
atches 494; Conservative 0; Mismatches 261; Indels 12; Gaps 4;

1 ATGCTTGGAAATGGATTGTTGTCATTCTTCTTCTTGGCAGATCAAGAGGAAAGATT 60

6 ATGCTTGAATGGATATTCGCACTATTTCTTCTGACAGCAGAGGCTCTGAAG-T 64

61 TCGTATGAAGTTAGGTGTTTCAAGATGGTTTACCATGCGACAGGACTTCTCAC 120

65 TCGTATGACAGGCTGGATGTTTTCACATGATATTCATGCTTGGAACTGCGAGAGA 124

121 GAGTTGGTGGTTTACCCCTGGTCTCCAGAGAGATAAACACTCGTTTCCCTCTCACT 180

125 CCAATTTATAGTTTACCCCTGGAGCCAGAGAGATAGGTACCAGTTCCCTTACACA 184

181 ATACACATCCCAATGCCATCAGAGATCAGTGGGTTAATTTCTTCACTATCCAGCC 240

185 AGAGAAATGGTAATAGCAATCAGGAGATATTCGAGTTAATCCCTCAAAATTTGGGAGC 244

241 TCATATTTTGGAAACAGACAGATCACCCTATCAATAGCTGGATGAAACAGATGGC 300

245 TCAACTTCAAAACAGCAGGAGACAGATTGTTGTTCAATGATTATAGATGAAGGA 304

301 AAA-----TGGCAGAGACATGTGCAATGTTTGTTCACAGCTGGAGATATAAATGC 354

305 GAAGAAGGCTGGACCTCAGATCTGTGCAAGAGGATGCTTACAGTGGAGATGTGAACGT 364

355 ATTAATTTAGATGATCAACGGTTTCAGGG---AATACATCCATGCTGTAAACAATCTC 411

365 ATTGCTGTAGCTGGAANAAGTCAAGGTGTGAGTACTCTCAGGCTTCAACACGTC 424

412 CGTGTGTTGGTCTGAGTGGCTTATTTTATGTGTTCTCATGAAATTTTGAATAT 471

425 CGTGTGTTAGTGTGAAATTCATATTTTATCAGCGTGTCTCGGGATCAATATTCATAC 484

472 TCCCTTCTAAGTGCACTGATTGGCCACAGCTTGGGAGCACACCTGGCTGGGGAAGCT 531

Db 485 TCTTCAGCCCATGTCCACATCATTTGTCACAGCTTGGAGCACACAGTGGCAGGAGGCT 544  
QY 532 GGGTCAAGGATACAGGCTTGGGAAGATAACTGGGTGGAGCCAGCTGGGCAATTTTTC 591  
Db 545 GCGAAGAGGAGGCGGAGGTTGGAAGATAACTGGAGTGGACCTGCCCAAGCTTATTTT 604  
QY 592 CACAACACTCCAAAGGAAGTCAGGCTAGAGCCCTCGGATGCCAACCTTTGTCAGCTTAT 651  
Db 605 CAGACACTCATATTGATGATCAGCTGGATAAAGCATGACAGATTGTCATGTTATC 664  
QY 652 CATCAAAATCAGCTCGCATCTCTTTTGGACTTGGTTGGAAACCATTTGATGCTTGTGT 711  
Db 665 CACACAGATACAGCTCCCATATAATCCCACTTAG--GTTTGAATGCCAAGCAATAGGA 722  
QY 712 CATCTTGACTTTTACCCCAATGGAGGAGACACATCCAGGATGTGA 758  
Db 723 CATCTTGACTTTTATCCAAACGGAGGAGTGGAAATCCAGGATGTGA 769

RESULT 11

BU417084

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

FEATURES

source

1. .795

/organism="Gallus gallus"

/mol\_type="RNA"

/strain="Layer and broiler"

/db\_xref="taxon:9031"

/clone="CHEST615120"

/sex="Male and female"

/dev\_stage="adult"

/lab\_host="DH10B"

/clone\_lib="CSEORBL07"

/note="Organ: pancreas; Vector: pBluescript II KS(+);

Site 1: EcoRI; Site 2: NotI; Modification of pBluescript

II KS(+) [Stratagene] vector to accommodate cDNA produced

with the T-trimmed protocol (Construction of

uni-directionally cloned cDNA libraries from messenger RNA

for improved 3' end DNA sequencing by Glenn Fu, et al.

U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with

NotI and EcoRI. Ligate in double stranded adaptor

containing BspI and BamHI sites

[5'ggccgcgtgcagcccgatccgagaaag]

[5'aattcttttttcggatccgggtgcagc]"

BASE COUNT 236 a 161 c 200 g 198 t

ORIGIN

Query Match 20.1%; Score 282.4; DB 13; Length 795;  
 it Local Similarity 64.6%; Pred. No. 1.4e-64;  
 ches 487; Conservative 0; Mismatches 256; Indels 11; Gaps 4;

14 GATTTGTTGATCTTTCTTTGGGACATCAAGAGGAAAGAGTTGCTATGAAGGT 73  
 |||||  
 2 GATATTGCGACTATTCTTCTGTGCACAGCAGAGGCTCTGAAGTTTGTATGACAGGC 61  
 |||||  
 74 TAGGTTGTTTCAAGATGTTTACCATGACAGGAGCTTTCTCAACAGAGTTGTTAGTT 133  
 |||||  
 62 TGGATGTTTACAGATGATATGCCATGCTGGAACATGACAGAGACCAATTTATAGT 121  
 |||||  
 134 TACCTGTTTCCAGAGAGATAAACAATCTGTTTCTCTCTACACTATCAACATCCCA 193  
 |||||  
 122 TACCTGAGAGCCAGAGAGATAGTACCACTCTCTTACACAGAGAGAAATGTA 181  
 |||||  
 194 ATGCTTATCAGAGATCAGTGGGTTAAATCTTCAACTATCCAGCTCATATTTGGAA 253  
 |||||  
 182 ATAGCAATCAGAGATATCTGCGTTTAACTCCCTCAACAATGGGAGCTCAAACTTCAAAA 241  
 |||||  
 254 CAGACAGATCACCCGTATCAACATAGCTGGATGGAGAAACAGATGGCAAA-----TGGC 307  
 |||||  
 242 CAAGCAGGAGACAAGATTGTTGTTTCAATGATTTATAGATGAAGGAGAGGCTGGA 301  
 |||||  
 308 AGAGACATGCTCAATGTTGCTACAGCTGGAAGATATAAATGCAATTAATTTAGATT 367  
 |||||  
 302 CTTAGATCTGTGCAAGAGAGATGCTTACAGTGGAGATGTGAACGTGTTGCTGTAGCT 361  
 |||||  
 368 GGATCAACGGTTCCAGG---GAATACATCATCTGTGTAACAATCTCGTGTGTTGGTG 424  
 |||||  
 362 GGAAGAGAGTGCAGGCTGCTAGTCTCTCAGCTTCAACAACGTCGTTGTTAGTG 421  
 |||||  
 425 CTGAGTGGCTTATTTTATGATGTTCTCATGAAAAAATTTGAATTAATTCCTTCAAG 484  
 |||||  
 422 CTGAAATTTGATATTTTATCAGCGTCTCGCGATCAATATTCATACTCTTCAGCCAACG 481  
 |||||  
 485 TGCATTTGATGCGCACAGCTGGGAGCACCTCGCTGGGAGCTGGGTCAAGGATAC 544  
 |||||  
 482 TCACATCATTTGCTACAGCTTGGAGCACAGTGGCAGGAGAGCTGGCAAGGAGGC 541  
 |||||  
 545 CAGGCTTGGAGAGATAAATCTGGGTTGGACCCAGCTGGGCAATTTTTCACACACTCAA 604  
 |||||  
 542 CAGGCTTGGAGATAAATCTGGGTTGGACCCCTGCGCAACCTTATTTTCAAGACACTCAA 601  
 |||||  
 605 AGAAGTCAAGCTAGACCCCTGGATGCGCACTTTGTTGAGCTTATTCATACAAATGCGAG 664  
 |||||  
 602 TTGAAGTCAAGCTGGATAAAGCGATGTCAGAGTTTGTGATGTTTATCCACACAGATACAG 661  
 |||||  
 665 CTGCTATCTCTTTGAGCTTGGTGTGTTGGAACCATTTGATGCTTGTGTTTCTTCACTTTT 724  
 |||||  
 662 CTCCATATTTCCCACTTAGTTTGGAA-TGGCCAGCAATAGGACATTTGAC-TTT 719  
 |||||  
 725 ACCCAATGGAGGAGAGCAGTCGCCAGGATGTA 758  
 |||||  
 720 ATCCAAACGAGGAGTGGAAATGCCAGGAGTGA 753  
 |||||

T 12  
 864

ITION  
 603671065F1 CSEORBL07 Gallus gallus cdna clone ChEST615a18 5', mRNA  
 sequence.

SIGN  
 BU416864

ON  
 BU416864.1 GI:25909535

RDS  
 EST.

E  
 Gallus gallus (chicken)

ANISM  
 Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 1 (bases 1 to 761)

ENCE  
 Boardman,P.E., Sanz-Bzquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,  
 Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 COMMENT

A Comprehensive Collection of Chicken cDNAs  
 Curr. Biol. 12 (22), 1965-1969 (2002)  
 2235534  
 12445392

Contact: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology (UMIST)  
 )  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.  
 Location/Qualifiers  
 1. .761  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="Layer and broiler"  
 /db\_xref="taxon:9031"  
 /clone="ChEST615a18"  
 /sex="Male and female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="CSEORBL07"  
 Note="Organ: Pancreas; Vector: pBluescript II KS(+);  
 Site 1: EcoRI; Site 2: NotI; Modification of pBluescript  
 II KS(+) [Stratagene] vector to accommodate cDNA produced  
 with the T-trimmed protocol (Construction of  
 uni-directionally cloned cDNA libraries from messenger RNA  
 for improved 3' end DNA sequencing by Glenn Fu, et al.  
 U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with  
 NotI and EcoRI. Ligate in double stranded adaptor  
 containing BspI and BamHI sites  
 15'gagcgtgtcgagccggatccggaagaaag!  
 [5'aatttttttcggtccgggtgcagc!]  
 222 a 158 c 187 g 194 t

BASE COUNT 222 a 158 c 187 g 194 t

ORIGIN

Query Match 19.8%; Score 277.6; DB 13; Length 761;  
 Best Local Similarity 64.8%; Pred. No. 2.7e-63;  
 Matches 478; Conservative 0; Mismatches 249; Indels 11; Gaps 4;

QY 1 ATGCTTGAATTTGATTTGCTTCTTCTTGGCACATCAAGAGGAAAGAGTT 60  
 |||||  
 Db 26 ATGCTTGAATTTGATTTGCTTCTTCTTGGCACAGGAGGCTCTGAAGTT 85  
 |||||  
 QY 61 TCTATGAAAGTTAGGTTTCAAGATGTTTACATGACAGGACTTTCTCAACA 120  
 |||||  
 Db 86 TCTATGACAGGCTGGATGTTTCAAGATGATTTCCATGCTTGGAACTGCAGAGAGA 145  
 |||||  
 QY 121 GAGTTGGTAGTTTACCTGCTTCCAGAGAGATAGTCCGTTTAAATTTTCAACTATCCAGGCC 240  
 |||||  
 Db 146 CCAATTTATAGTTTACCTGAGCCAGAGAGATAGTCCAGTTCTTCTTACACACA 205  
 |||||  
 QY 181 ATACACAATCCCAATGCCCTATCAGGAGATCAGTCCGTTTAAATTTTCAACTATCCAGGCC 240  
 |||||  
 Db 206 AGAGAAATGTTATAGCAATCAGGAGATATCTCAGTTAATCCCTCAACAATTTGGGAGC 265  
 |||||  
 QY 241 TCATATTTTGGACAGACAGATCACCCGTTATCAACATAGCTGATGGAAGAGAGATGAGGC 300  
 |||||  
 Db 266 TCAAACTTCAAAACAGCAGGAGAGAGATTTGTTTTCATGATTTATAGATGAAGGA 325  
 |||||  
 QY 301 AAA-----TGGCAGAGAGACATGTGCAATGTTTGTGTACAGCTGGAGAGATATAAATGC 354  
 |||||  
 Db 326 GAAGAGGCTGGACCTCAGATCTGTGCAAGAGGATGCTTACAGTGGAGAGATGTAAGTGT 385  
 |||||  
 QY 355 ATTAATTTAGATTTGATCAAGGTTTCAAGG---GAATACATCCATGCTGTAAACAATCTC 411  
 |||||  
 Db 386 ATTGCTGAGACTGGAAGAGAGGTCAGAGGTGTCACTACTCTCAGGTTTCAACAACAGCTC 445  
 |||||  
 QY 412 CGTGTGTTGTTGCTGAGGCTGCTTATTTTATTCATGTTCTCATGAAAAAATTTGAATAT 471  
 |||||  
 Db 446 CGTGTGTTGAGGTTGCTGAAATTCATATTTTATCAGGCTGCTCGCGGATCATATATTCATAC 505  
 |||||



Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A., and Hubbard, S.J.  
 A Comprehensive Collection of Chicken cDNAs  
 Curr. Biol. 12 (22), 1965-1969 (2002)  
 22335534  
 JMED 12445392  
 Contact: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology (UMIST)  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.  
 Location/Qualifiers  
 1. .781  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="Layer and broiler"  
 /db\_xref="taxon:9031"  
 /clone="CHES7616m22"  
 /sex="Male and female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="CSEORBL07"  
 /note="Organ: pancreas; Vector: pBluescript II KS(+);  
 Site 1: EcoRI, Site 2: NotI; Modification of pBluescript  
 II KS(+) [Stratagene] vector to accommodate cDNA produced  
 with the T-trimmed protocol (Construction of  
 uni-directionally cloned cDNA libraries from messenger RNA  
 for improved 3' end DNA sequencing by Glenn Fu, et al.  
 U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with  
 NotI and EcoRI. Ligate in double stranded adaptor  
 containing BspI and BamHI sites  
 [5'ggccgctgagcccgatccgagcgaagaaag]  
 [5'aattcttttttcgagtcgggctgcagc]"  
 233 a 159 c 195 g 194 t

RBS  
 source

COUNT 233 a 159 c 195 g 194 t  
 N  
 ry Match 19 5%; Score 274; DB 13; Length 781;  
 t Local Similarity 65.1%; Pred. No. 2.5e-62;  
 ches 500; Conservative 0; Mismatches 255; Indels 13; Gaps 6;  
 1 ATGCTTGGAAATTTGGATTTGCTTCTTCTTTGGCAGATCAAGAGGAAAGAGATT 60  
 8 ATGCTTGGAAATTTGGATTTGCTTCTTCTTGGCAGATCAAGAGGAGGCTCTGAATT 67  
 61 TGTATGAAGATTTAGGTTTCAAGATGTTTACCATGAGCAGAGATTTCTCAACA 120  
 68 TGTATGACAGCTGGGATTTTCAAGATGATATGCCATGCTGTGGAAGTGCAGAGAGA 127  
 121 GAGTTGGTGGTTTACCTGGTCTCCAGAGAGATAACACTGCTTTCTGCTCTACACT 180  
 128 CCAATTTATAGTTTACCTGGAGCCAGAGAGATAGTAGTCCAGTTCTCTTACACACA 187  
 181 ATACAAATCCCAATGCCCTATCAGAGATCAGTGGGTTAATTTCTCAACTATCAAGGCC 240  
 188 ACAGAAATGGTAATAGCAATCAGAGATATCTGCAGTTAATCCTCAACAATTGGGAGC 247  
 241 TCATATTTTGGACACAGATATCCCGTATCAACATAGCTGATGGAAGAACAGATGGC 300  
 248 TCAACTTCAAAACAGAGAGGAGAGAGATTTGTTGTATGATTTATAGATGAAGGA 307  
 301 AAA-----TGGCAGAGAGACATGTGCAATGTGTTGTCTACAGCTGGAGAGATATAAATTGC 354  
 308 GAAGAGGCTGGACCTCAGATCTGTGCAAGAGGATGCTTACAGTGGAGATGTGAATGT 367  
 355 ATTAATTTAGATGATCAAGCTTACAGGG---AATACATCCATGCTGTAACAATCTC 411  
 368 ATGCTGTAGACTGGAAAAAGGTGCAAGGTTGTCAGTACTCTCAGGCTTCAAAACAGCTC 427  
 412 CTTGTTGTTGGTGTCTGAGTGGCTTATTTTATGATGTTCTCATGAAAAAATTTGATAT 471  
 428 CTTGTTGTTAGTGTCTGAAATTTGCATATTTTATCAGCTGCTCGCGGATCAATATTCATC 487

QY 472 TCCCTTCTAAAGTCCACTTGTATGGCCACACAGCTTGGAGCACACCTGGCTGGGAGAGCT 531  
 Db 488 TCTTACGCCACGCTCCATCATTTGTTCACAGCTTGGAGCACACCTGGAGGAGAGGCT 547  
 QY 532 GGGTCAAGGATACCAAGGCTTGGAGAGATAAATGGTTGGACCCAGCTGGGCCCAATTTTC 591  
 Db 548 GGAAGAGAGGAGGAGGAGG-GTTGGAAGAGATAAATGGAGCTGGAGCCCTGC-CCAACTTATTTT 605  
 QY 592 CACAACACTCCAAAGGAGTCAAGCTAGACCCCTGGATGCGAATGCAACTTTGTTGAGCTTAT 651  
 Db 606 CAAGACACTCCAAATTTGAAGTCAAGCTAGATTAAGACCGATGAGAGTTTGTGTGATGTTATC 665  
 QY 652 CATACAAATGCAAGCTCGCATCTCTTTGAGCTTGGTGTGGAACCAATGATGCTGTGGGT 711  
 Db 666 CACACAGATACAGCTCCCATAT-TCCCACTTAGTTTGGAAATGGCCCAAGCAATAGGA 724  
 QY 712 CATCTTGAATTTTACCAAAATGGAGG-GAAGCAGATGCGAGGATGTA 758  
 Db 725 CATCTTGAATTTTATCCAAACGAGGAGTGTGAAATGCCAGGTTGTA 772  
 RESULT 15  
 BU416394  
 LOCUS  
 DEFINITION 741 bp mRNA linear EST 29-NOV-2002  
 603570888F1 CSEORBL07 Gallus gallus cDNA clone CHES7614p21 5', mRNA  
 sequence.  
 BU416394  
 ACCESSION  
 VERSION BU416394.1 GI:25909065  
 KEYWORDS  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 REFERENCE 1 (bases 1 to 741)  
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,  
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.  
 A Comprehensive Collection of Chicken cDNAs  
 Curr. Biol. 12 (22), 1965-1969 (2002)  
 22335534  
 PUBMED 12445392  
 CONTACT: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology (UMIST)  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.  
 Location/Qualifiers  
 1. .741  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="Layer and broiler"  
 /db\_xref="taxon:9031"  
 /clone="CHES7614p21"  
 /sex="Male and female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="CSEORBL07"  
 /note="Organ: pancreas; Vector: pBluescript II KS(+);  
 Site 1: EcoRI, Site 2: NotI; Modification of pBluescript  
 II KS(+) [Stratagene] vector to accommodate cDNA produced  
 with the T-trimmed protocol (Construction of  
 uni-directionally cloned cDNA libraries from messenger RNA  
 for improved 3' end DNA sequencing by Glenn Fu, et al.  
 U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with  
 NotI and EcoRI. Ligate in double stranded adaptor  
 containing BspI and BamHI sites  
 [5'ggccgctgagcccgatccgagcgaagaaag]  
 [5'aattcttttttcgagtcgggctgcagc]"  
 217 a 154 c 183 g 187 t  
 BASE COUNT

FEATURES  
 source

IN

```
ery Match          19.4%; Score 271.8; DB 13; Length 741;
st Local Similarity 64.3%; Pred No. 9.5e-62;
tches 474; Conservative 0; Mismatches 252; Indels 11; Gaps 4;

1  ATGCTTGAAATTTGGATTGTTGCAATCTTCTTTGGCCACATCAAGAGGAAAGAAAGTT 60
   |||||
7  ATGCTTCGAATTTGGGATATTGCGCTATTCTCTCTGACAGCGAGGCTCTGAAGTT 66

61  TGCATATGAAGTTAGGGTGTTCARAGATGGTTTACCATGGACCGAGGACTTCTCAACA 120
   |||||
67  TGCTATGACAGCTGGGATGTTTACAGATGATATGCCATGGTCTGGAATCGCAGAGAGA 126

121  GAGTTGGTGTGTTTACCTGTCTCTCCAGAGAAGATAACACTCGTTTCTGCTCTACACT 180
   |||||
127  CCAATTTATAGTTTACCTGGAGCCACAGAGATAGGTACCCAGTTCTCTTACACACA 186

181  ATACACATCCCAATGCCCTATCAGAGATCAGTGGGTTAAATCTTCAACTATCCAGCC 240
   |||||
187  AGAGAAATGTGTAATAGCAATCAGGAGATATCTGCAGTTAATCCCTCAACAATTGGGAGC 246

241  TCATATTTTGGAAACAGACAGATCACCCGTATCAACATAGCTGATGGTGAAGAACAGATGCC 300
   |||||
247  TCARACTTCAAAACAGCAGAGACAGATTTGTTGTTTCATGGATTTATAGATGAGAGA 306

301  AAA-----TGGCAGAGAGACATGTGCAATGTGTGCTACAGCTGGAAGATATAAATTGC 354
   |||||
307  GAAGAAGGCTGGACCTCAGATCTGTGCAAGAGGATGCTTACAGTGGAGATGTGAACGTGT 366

355  ATTAATTTAGATTTGGATCAAGGTTTACGG--GATACATCCATGCTGTAAACATCTC 411
   |||||
367  ATTGCTGTAGACTGGAAAAAGGTGCAAGGTGTCAGTCTCTCAGGCTTCGAACACGTC 426

412  CGTGTGTGTGTGCTGAGGTGGCTTATTTATTGATGTTCTCATGAAAAAATTTGAATAT 471
   |||||
427  CGTGTGTGTGCTGAAATGTCATATTTTATCAGCGTGTCTCGCGGATCAATATTCTATAC 486

472  TCCCTTCTTAAAGTGCACCTTGATTGGCCACAGCTTGGAGACACACCTGCTGGGAGCT 531
   |||||
487  TCTTCAGCCAATGTCCACATCAITGGTCACAGCCTTGGAGCACACGTGGCAGGAGGCT 546

532  GGGTCAGAGTACCAGGCTTGGAAAGATAACTGGGTGGACCCAGCTGGGCCATTTTTC 591
   |||||
547  GGCAAGAGGAGGCCAGG-GTTGGAAGATAAATGGACTGGACCTGCCCAACCTTATTTT 605

552  CACAACACTCCAAAGGAAGTCAGGCTAGACCCCTCGGATGCCCACTTTGTTGACGTTAT 651
   |||||
606  CAAGACACTCCAATTTGAAGTCAGGCTGGATAAAAGCGATGCAGAGTTTGTGATGTTATC 665

652  CATACAAATGCAGCTCGCATCTCTTTGAGCTTGGTGTGGAAACCATTTGATGTTGTGGT 711
   |||||
666  CACACAGATACAGCTCCCATAT-TCCCACTTAGGTTCGGAATGSCCCCAAGCAATAGGA 724

712  CATCTTGACTTTTATCCC 728
   |||||
725  CATCTTGACTTTTATCC 741
```

ch completed: January 6, 2004, 11:10:44  
time : 3198 secs